INVESTIGATION OF THE BIOSYNTHESIS OF THE PHOSPHONATE ANTIBIOTICS BIALAPHOS AND DEHYDROPHOS

BY

JIN HEE LEE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Chemistry
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor Wilfred A. van der Donk, Chair
Professor Steven C. Zimmerman
Associate Professor Scott K. Silverman
Professor Neil L. Kelleher

ABSTRACT

Phosphonate and phosphinate natural products possess a range of biological activities as a result of their ability to mimic phosphate esters or tetrahedral intermediates formed in enzymatic reactions involved in carboxyl group metabolism. The widespread use in medicine and agriculture of these compounds has resulted in active researches, revealing novel biochemistry behind their metabolic pathways. One of such compounds, PTT, contains unusual amino acid phosphinothricin attached to two alanine residues. The peptide bond formation and methylation on phosphorus in PTT are key steps in PTT biosynthesis, however, the exact timing of these transformations was not known. To provide insights into these questions, a heterologous expression system was developed for PhsA, the first of three NRPS proteins involved in PTT biosynthesis, and k_{cat}/K_m value for the ATP-pyrophosphate exchange activities of the purified protein with putative substrates were determined. The k_{cat}/K_m value for L-*N*-AcPT was 7-fold higher than for D,L-*N*-AcDMPT, suggesting the former might be the physiological substrate. Both substrates were loaded onto the thiolation domain of PhsA as demonstrated by FTMS. The possible implications of these findings for the timing of the methylation reaction are discussed. Another phosphonate presented here is antibiotic dehydrophos, which contains a unique methyl ester of a phosphonodehydroalanine group. Since the dianionic form of phosphonates at pH 7 poses a drawback with respect to their ability to mimic carboxylates and tetrahedral intermediates, esterification of a phosphonate group by SAM-dependent methyltransferase, DhpI would provide a solution to alleviating the high charge state, and thus improve the bioavailability of phosphonates. This study describes expression, purification and reconstitution of DhpI *in vitro*. Also, substrate scope and X-

ray crystal structure of DhpI are also discussed. The enzyme utilizes *S*-adenosylmethionine to methylate a variety of phosphonates including 1-HEP, 1,2-DHEP, and Ac-1-AEP. Kinetic analysis showed that the best substrates are tripeptides containing as C-terminal residue a phosphonate analog of alanine suggesting the enzyme acts late in the biosynthesis of dehydrophos. These conclusions are corroborated by the X-ray structure that reveals an active site that can accommodate a tripeptide substrate. Furthermore, the structural studies demonstrate a novel conformational change brought about by substrate or product binding. Interestingly, the enzyme has low substrate specificity and was used to methylate the clinical antibiotic fosfomycin and the clinical anti-malaria candidate fosmidomycin, showing its promise for applications in bioengineering. Overall, the findings in this thesis provide valuable insights in Nature's strategy for making novel phosphonates which can be further developed into applications in novel drug discovery.

To my family

ACKNOWLEDGEMENTS

There are many people without whom I could have not accomplished this work. I am grateful to have known them as my teachers, coworkers, collaborators, and friends who have helped me in many various ways during my graduate study.

I would like to thank Professor Wilfred van der Donk for giving me the opportunity to work on this exciting research, for his guidance and patience, and for being a great role model as a scientist and a teacher. I thank Professor Steven Zimmerman, Professor Neil Kelleher, and Professor Scott Silverman for serving on my thesis committee and for their advice and encouragement. Also, I thank Professor Bill Metcalf, Satish Nair and Neil Kelleher who have guided me develop this multidisciplinary project into a success by providing their insights, helpful discussions and suggestions. I was also fortunate to collaborate with very talented researchers, Dr. Brian Bae, Benjamin Circello, and Brad Evans who have contributed many of works which are in this thesis. It was a great, eye-opening experience to work with them from such different scientific background.

Members of the van der Donk and MMG groups have been a great source of help, encouragement and fun. I thank the phosphonate subgroup members, especially Dr. Robert Cicchillo for his great help and teaching me learn many important experimental techniques and Dr. Svetlana Borisova for her encouraging atmosphere she brings in to the lab. I also thank Dr. Michael Kuemin and Dr. Gongyong Li for providing synthetic compounds without which this work could have not been completed, and undergraduate student Amanda Brunner whose assistance with DhpI project was a great help. I am glad

to have known my labmates, Dr. Leigh Anne Furgerson, Juan Velasquez, Joel Cioni, and Dr. Heather Cooke for the laughs and encouragements during my times as a graduate student.

I thank Michael Vu from the Martinis group and Dana Baum from the Silverman group who guided me to set up the radioactivity experiment in my lab. I would like to acknowledge Prof. Dr. Wolfgang Wohlleben and Dr. Ullrich Keller at the University of Tübingen, Germany for sharing their unpublished data on PhsA study.

All the friends who I have made outside the research have been important in helping me focus on my career goal and maintain perspectives during my time at UIUC. The times and stories shared together with Professor Sua Myong, Dr. Heekyung Kim, Hyunju Shin, and the members of Isaiah fellowship will be cherished for the rest of my life. Foremost, I would like to thank my family for their support and prayers. My mother who always believed in me, my sister who is always so giving and sacrificing, my brother who has travelled all the way from Mexico just to be with me for my graduation ceremony, my amazing two children, Eun Seo and Joshua, for being the greatest source of my motivation for what I am doing. Finally, I am grateful to have shared this entire journey with my husband Hyun Soo Lee and would like to thank him for his endless love, patience, and support. Without all these helps, this work would have not been possible.

.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

LIST OF ABBREVIATIONS

A domain	Adenylation domain
Ac-1-AEP	1-Acetylaminoethyl phosphonic acid
AcCoA	Acetyl coenzyme A
AcDMPT	*N*-acetyldemethylphosphinothricin
AcPT	*N*-acetylphosphinothricin
AdoHcy	*See* SAH
AdoMet	*See* SAM
AEP	Aminoethylphosphonate
AlaP	Phosphonoalanine
APPA	Aminophosphonopentenoic acid
ATP	Adenosine triphosphate
B. subtilis	*Bacillus subtilis*
BSA	Bovine serum albumin
C domain	Condensation domain
1,2-DHEP	1,2-Dihydroxyethylphosphonate
1,2-DHEP-OMe	1,2-Dihydroxyethylphosphonate methyl ester
DTT	Dithiothreitol
E. coli	*Escherichia coli*
E domain	Epimerization domain
FPLC	Fast protein liquid chromatography
FTMS	Fourier-Transform mass spectrometry

1-HEP	1-Hydroxyethylphosphonate
2-HEP	2-Hydroxyethylphosphonate
1-HEP-OMe	1-Hydroxyethylphosphonate methyl ester
2-HEP-OMe	2-Hydroxyethylphosphonate methyl ester
IPTG	Isopropyl-*β*-D-thiogalactopyranoside
IS	Internal standard
LB	Luria-Bertani
MBP	Maltose binding protein
N-Met	*N*-Methylation
NRPS	Nonribosomal peptide synthetases
NS	Nozzle-Skimmer
PCP	Peptidyl carrier protein
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PepM	Phosphoenolpyruvate mutase
PnAA	Phosphonoacetaldehyde
PnPy	Phosphonopyruvate
PPant	Phosphopantetheine
PP_i	Pyrophosphate
PT	Phosphinothricin
PTT	Phosphinothricin tripeptide
RPLC	Reverse phase liquid chromatography
RT	Retention time

S. hygroscopicus	*Streptomyces hygroscopicus*
S. luridus	*Streptomyces luridus*
S. viridochromogenes	*Streptomyces viridochromogenes*
SAH	*S*-adenosyl-L-homocysteine
SAM	*S*-adenosyl-L-methionine
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TCEP	Tris(2-carboxyethyl)phosphine
T domain	Thiolation domain
Te domain	Thioesterase domain

CHAPTER 1: INTRODUCTION

1.1 – Antibiotics Background

Antibiotics have been enjoying the reputation of being the most successful chemotherapy developed in the twentieth century. Natural products with antimicrobial activity are produced by both fungi and bacteria as secondary metabolites used to combat foreign microbes.[1] Exploiting natural products led to the discovery of many new antibiotics during the 1940s-1960s, followed by extensive interrogations of these novel chemical scaffolds to tailor them into drugs with improved pharmacokinetics.[2] The major targets of current commercial antibiotic drugs are cell wall biosynthesis, protein biosynthesis, and DNA gyrases (Table 1.1).[1,3]

The widespread use of antibiotics in the last century has resulted in the emergence of antibiotic resistance (Table 1.1). Antibiotic resistance is almost inevitable as there is a continuous evolutionary pressure on microorganisms which introduce genetic mutations at random; selection for resistant mutants in the population results in resistance. Ever since the advent of antibiotics for the treatment of bacterial infections, antibiotic resistance has been a continuing problem and the search for new antibiotics has been of primary importance.[4] Multidrug-resistant bacteria that cannot be treated with currently available drugs are emerging at an alarming rate, posing a threat to mankind. In fact, infectious diseases are the major causes of death worldwide.[5] Yet industrial efforts to discover novel antibiotics have declined markedly. Causes for decreased support of antibiotics research, among others, include lack of industry productivity, increasing size

of clinical trials, low projected profits for non-chronic diseases, and increased generic competition.[6]

In order to overcome the imminent threat of antibiotic resistance, new approaches for antibiotic discovery must be sought, that embrace modern chemistry, bioinformatics, and biological tools. As shown in Table 1.1, most of the antibiotics that are currently in clinical use are derived from natural products discovered during the heydays of natural product discovery. These compounds are usually produced by soil bacteria (e.g. *Streptomyces*) that can be cultured in a laboratory setting.[5] Screening for strains capable of producing biological activities, optimization of culturing conditions for antibiotic production, isolation of the natural product responsible for the biological activity, and subsequent structural identification and synthesis of the complex natural products are highly demanding processes. Accordingly, modern drug discovery has turned away from the labor-intensive and costly natural product research to target-based combinatorial chemical approaches. These changes, brought by the advances in high-throughput screening and ever-increasing genomic information, also provided new opportunities for revolutionizing antibiotic mining strategies. At present, the pursuit of novel biological targets by high-throughput-screening approaches has not been very successful so far, and still needs to overcome many hurdles.[4] These challenges are also reflected in the renewed interest in natural product antibiotics research in recent years, which emphasizes finding new targets using reverse-genomics as well as revisiting old compounds for further investigation. This thesis presents efficient methodology by which a class of naturally produced antibiotics, phosphonates and phosphinates, can be discovered and their biosynthesis investigated by reconstitution of the enzymatic reactions.

Class of Antibiotics	Derivatives	Bacterial Target	Year Deployed	Resistance Observed
Sulfonamides	Sulfanilamide Sulfadiazine	Fatty acid metabolism	1930s	1940s
Quinolones	Ciprofloxacin Ofloxacin	DNA replication (gyrase)	1960s	1980s
Oxazolidones	Cycloserine Linezolid	Protein synthesis	2000	2001
Penicillins	Amoxicillin Ampicillin	Cell wall biosynthesis	1943	1946
Aminoglycosides	Gentamicin Streptomycin	Protein synthesis	1943	1959
Tetracyclines	Tigecycline Minocycline	Protein synthesis	1948	1959
Macrolides	Clarithromycin Erythromycin	Protein synthesis	1952	1988
Glycopeptides	Vancomycin Telavancin	Cell wall biosynthesis	1956	1988
Cephalosporins	Cefaclor Cefuroxime	Cell wall biosynthesis	1960s	1960s
Lipopeptide	Daptomycin Surfactin	Cell membrane	2003	2004

Table 1.1 Different classes of antibiotics and their derivatives. The shaded antibiotics are derived from synthetic lead compounds while the other antibiotics are derived from natural products.[1,3,7]

1.2 – Phosphonates and Phosphinates

Phosphonic and phosphinic acids (Figure 1.1) contain carbon-phosphorus (C-P) bonds, and are known to exert important biological activities as they are structural analogues of naturally prevalent molecules such as carboxylic acids and phosphate esters and anhydrides. Some known phosphonate and phosphinate compounds isolated from Nature have proven to act as potent competitive inhibitors against many enzymes. Also when chosen appropriately the inhibitory activities of these compounds generally do not affect enzymatic processes taking place in humans, making them attractive drug candidates. Despite the demonstrated utility of phosphonates as therapeutics and agricultural chemicals,[8] until very recently our knowledge of their biosynthesis had been very limited. Recent progress in elucidating biogenesis of C-P compounds in a collaborative effort named "Mining Microbial Genomes for Novel Antibiotics" at the Institute for Genomic Biology of University of Illinois has resulted in the discovery of many important metabolic pathways; moreover it has shown that there are many C-P natural products yet to be discovered.

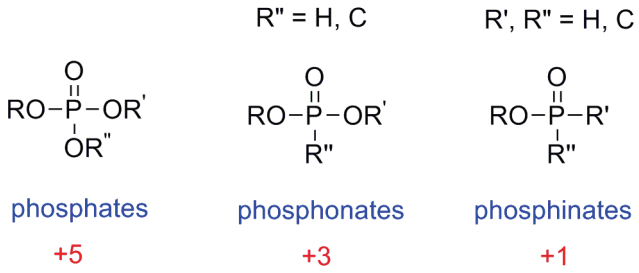


Figure 1.1 Chemical structure and formal oxidation state of compounds containing phosphorus. Phosphonates contain one bond of phosphorus to carbon or hydrogen, and phosphinates contain two bonds between phosphorus and carbon or hydrogen. Phosphates have the highest oxidation state of +5.

While phosphate esters are unstable, phosphonates and phosphinates are highly resistant to hydrolysis and enzymatic cleavage. These characteristics confer phosphonates and phosphinates the ability to inhibit a wide variety of cellular transformations adaptable to medicinal and agricultural applications. Notable examples of these compounds include aminophosphonopentenoic acid (APPA), a strong antifungal and antibacterial agent which inhibits threonine synthesis;[9] fosfomycin, an FDA-approved antibiotic;[10] phosphinothricin tripeptide (PTT), a widely used herbicide;[11] FR-900098 and fosmidomycin, which inhibit DXP reductioisomerase;[12,13] and K-26, a strong inhibitor of angiotensin-converting enzyme (Figure 1.2).[14] Many efforts have focused in recent years to isolate the biosynthetic gene clusters of these compounds and to characterize their biosynthetic pathways,[8] which resulted in partial elucidation of biosynthetic steps involved in production of some biochemically and therapeutically important C-P containing compounds shown in Figure 1.3. Although naturally occurring C-P compounds are also found in larger phosphonate-containing polymers with structural functions in the form of phospholipids, polysaccharides, and proteins, the focus of this thesis will be confined to small molecule C-P compounds exhibiting antibiotic activities.

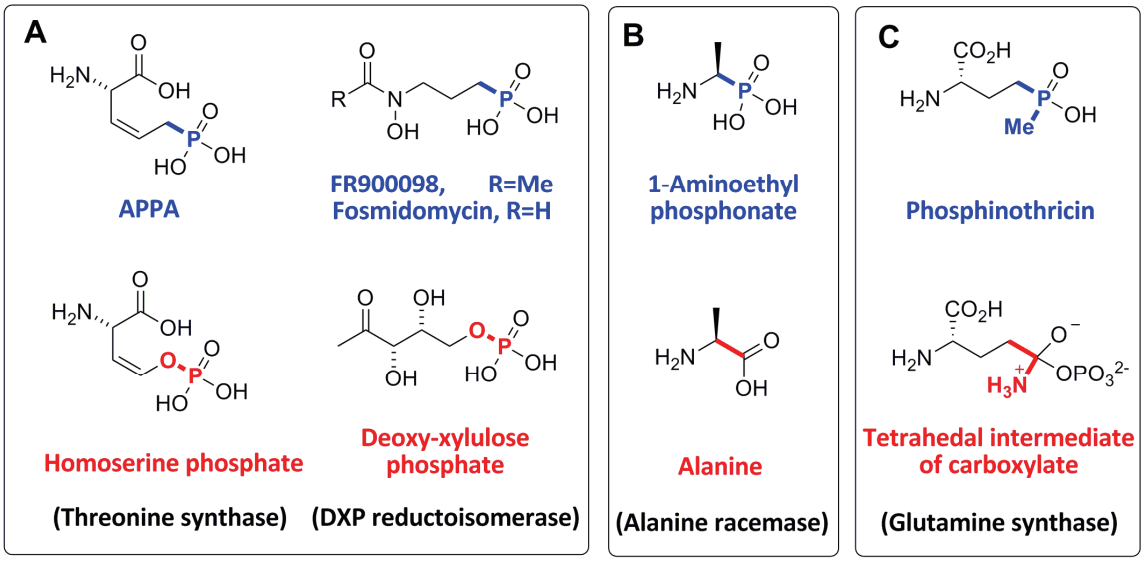


Figure 1.2 Examples of natural phosphonic and phosphinic acids. These C-P bond containing compounds (blue) can inhibit target enzymes (in parentheses, black) as structural mimics of enzyme substrates and intermediates (red). Shown here are (A) mimics of phosphate esters (APPA, FR900098 and fosmidomycin), (B) a mimic of carboxylate group (1-aminoethyl phosphonate), and (C) a mimic of the tetrahedral intermediate generated by ammonia addition during glutamine synthetase activity (phosphinothricin).

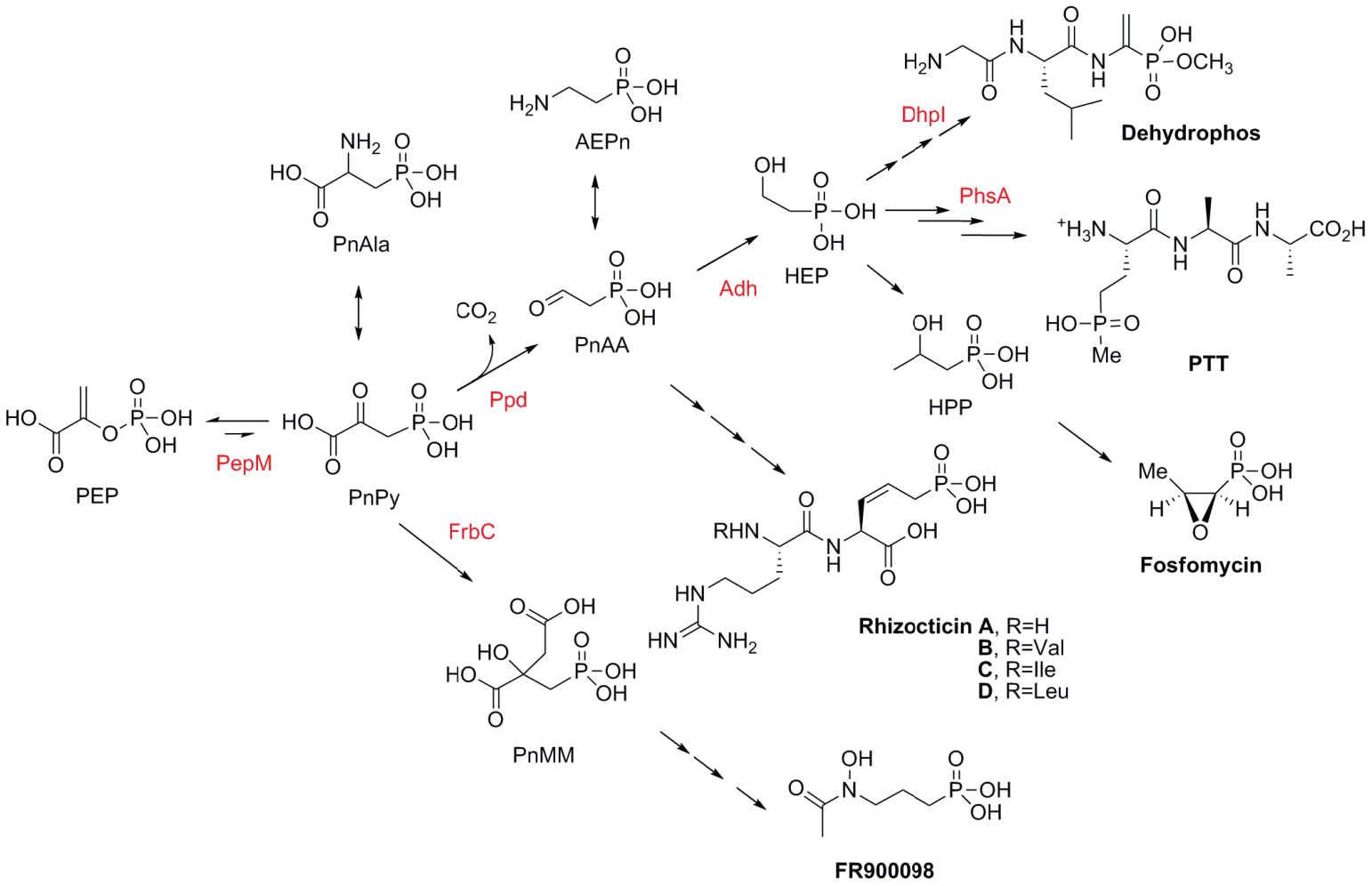


Figure 1.3 Overview of phosphonate and phosphinate biosynthesis. Biosyntheses of many important phosphonate compounds share the same first step of isomerization of phosphoenolpyruvate (PEP) to phosphonopyruvate (PnPy) by PepM (Drawing modified from *Annual Review of Biochemistry* **2009**, *78*, 65-94).

The increased available information on the biosynthesis of C-P compounds in recent years[8] demonstrated that the phosphorus bond formation in phosphonate and phosphinate biosynthesis is accomplished by the same initial step catalyzed by phosphoenolpyruvate (PEP) mutase (Figure 1.3).[8] This reaction is thermodynamically unfavorable and is often coupled to a subsequent decarboxylation reaction catalyzed by phosphonopyruvate (PnPy) decarboxylase (Figure 1.3).[8] The reliance of the biosynthetic pathways of C-P compounds on PEP mutases (PepM) has been used to screen and identify gene clusters for phosphonate production. Briefly, degenerate primers for *pepM* were designed for polymerase chain reaction (PCR), which can successfully identify phosphonate biosynthetic gene clusters from the genomic DNA of a known producing strain and even from uncharacterized environmental samples.[8] In fact, gene clusters of several prominent C-P compounds such as PTT,[15] dehydrophos,[16] fosfomycin,[10] and FR900098[12] were identified in producing strains using this approach.

Understanding the strategies that Nature employs to make novel phosphonates not only gives an insight in their biogenesis but is a key to the improvement in production and potential downstream combinatorial biosynthesis. Also, given the broad biological activities for the currently known phosphonates reported, investigation of unexplored phosphonates would pave a way to discovery of novel compounds with high potency. The exotic functional groups present in these compounds can serve as inspirations for synthetic chemists designing novel chemical scaffolds and also for finding new modes of action. Furthermore, their biosynthetic pathways include a number of unprecedented enzymatic reactions worth investigating, and these enzymes may be developed into biocatalysts.

1.3 - Nonribosomal Peptide Synthesis

Natural product antibiotics are often synthesized as short peptides by multienzyme complexes called nonribosomal peptide synthetases (NRPSs) instead of being synthesized on the ribosome. Because they are not limited by tRNA synthetases, nonproteinogenic amino acids can be incorporated into a peptide product, and various tailoring reactions can take place to give rise to a peptide decorated with various functional groups.[17] Examples of peptide antibiotics synthesized by NRPSs include tyrocidine, gramicidin S, bacitracin, vancomycin, and daptomycin (Figure 1.4).[17] The lengths of these peptides can vary from 2 to 20 amino acids, and can be produced as macrocyclic or linear peptides.[18]

NRPS assembly lines consist of multi-domain enzymes organized into modules. Each module is responsible for incorporating one amino acid (sometimes several) building blocks into the growing peptide chain. Generally there is a chain initiation module at the *N*-terminus, followed by elongation modules, and finally a termination module at the *C*-terminus (Figure 1.5). In elongation modules, minimal compositions are condensation (C), adenylation (A), and thiolation (T) domains, the latter also known as peptidyl carrier protein (PCP). An A domain is responsible for recognizing and activating the amino acid substrate, and therefore acts as a gate keeper in selecting very specific amino acids. The amino acid is activated as aminoacyl adenylate in the presence of ATP and Mg^{2+} which resembles the activation process for tRNA synthetase in ribosomal peptide synthesis. Prior to loading the activated aminoacyl adenylate, the T domain undergoes posttranslational modification known as phosphopantetheinylation, a priming process by which the 4'-phosphopantetheine (PPant) group from acetyl CoA is

transferred by a PPant transferase to a highly conserved active site serine residue in the T domain (Figure 1.5). This PPant group is a handle-like long chain structure with a terminal thiol group which covalently tethers the aminoacyl adenylate to the T domain via a thioester bond. The PPant arm is also required for flexible positioning of the amino acid attached to the T domain of one module to near the C-domain of the next module, after which thiotemplate-guided peptide bond formation occurs in the C domain. After the elongation of the entire peptide is completed, the final product is released from the enzyme complex by a thioesterase (Te) domain. Modifications to amino acids can take place by various reactions such as epimerization (E) and *N*-methylation (*N*-Met) while still bound to the enzymes by tailoring domains in NRPSs or they can be achieved after the peptide is released from the synthetase.

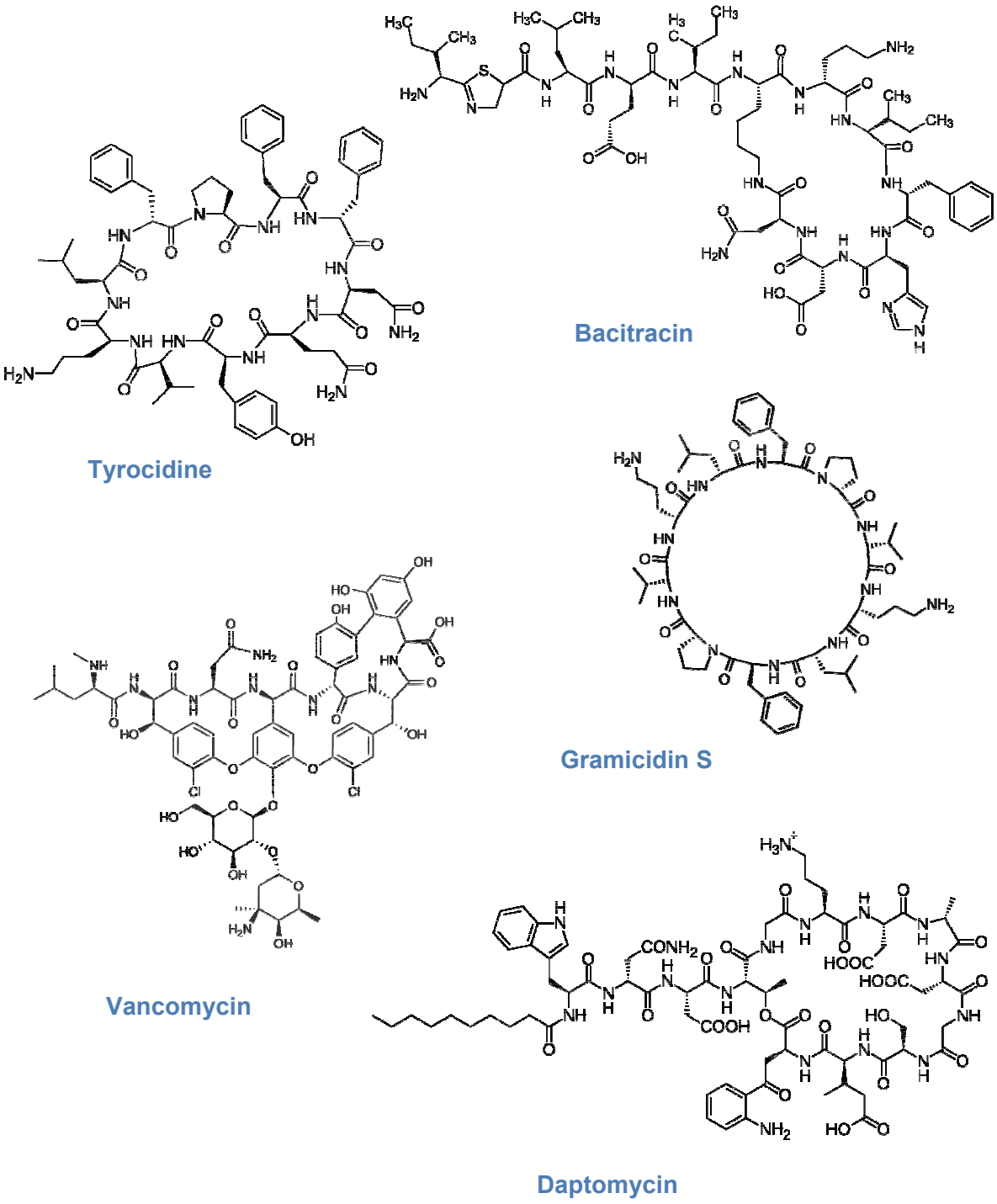


Figure 1.4 Examples of peptide antibiotics assembled by nonribosomal peptide synthetases (NRPSs).

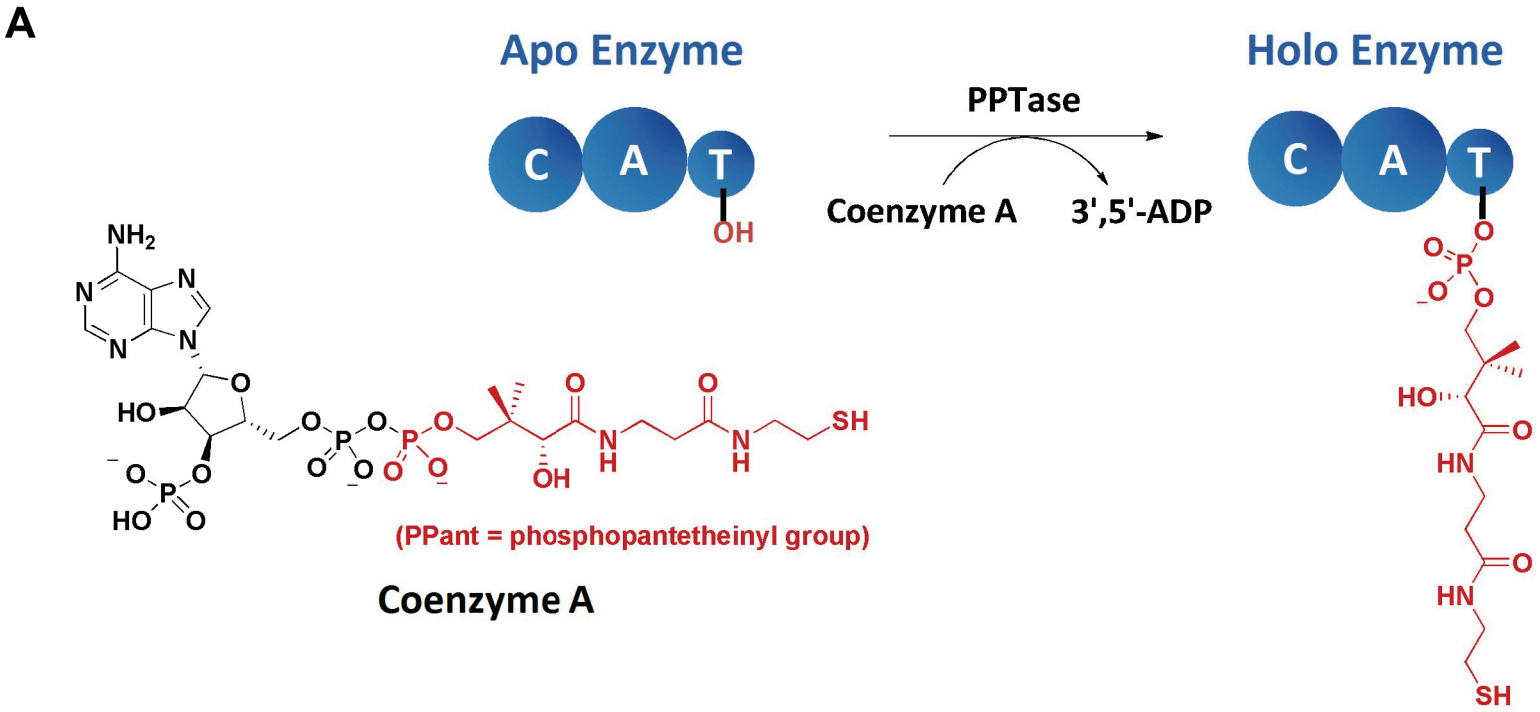


Figure 1.5 General organization of NRPS assembly line showing (A) conversion of an apo T to a holo T domain by posttranslational phosphopantetheinylation, (B) three modules with multiple domains, and (C) reactions taking place in core domains.

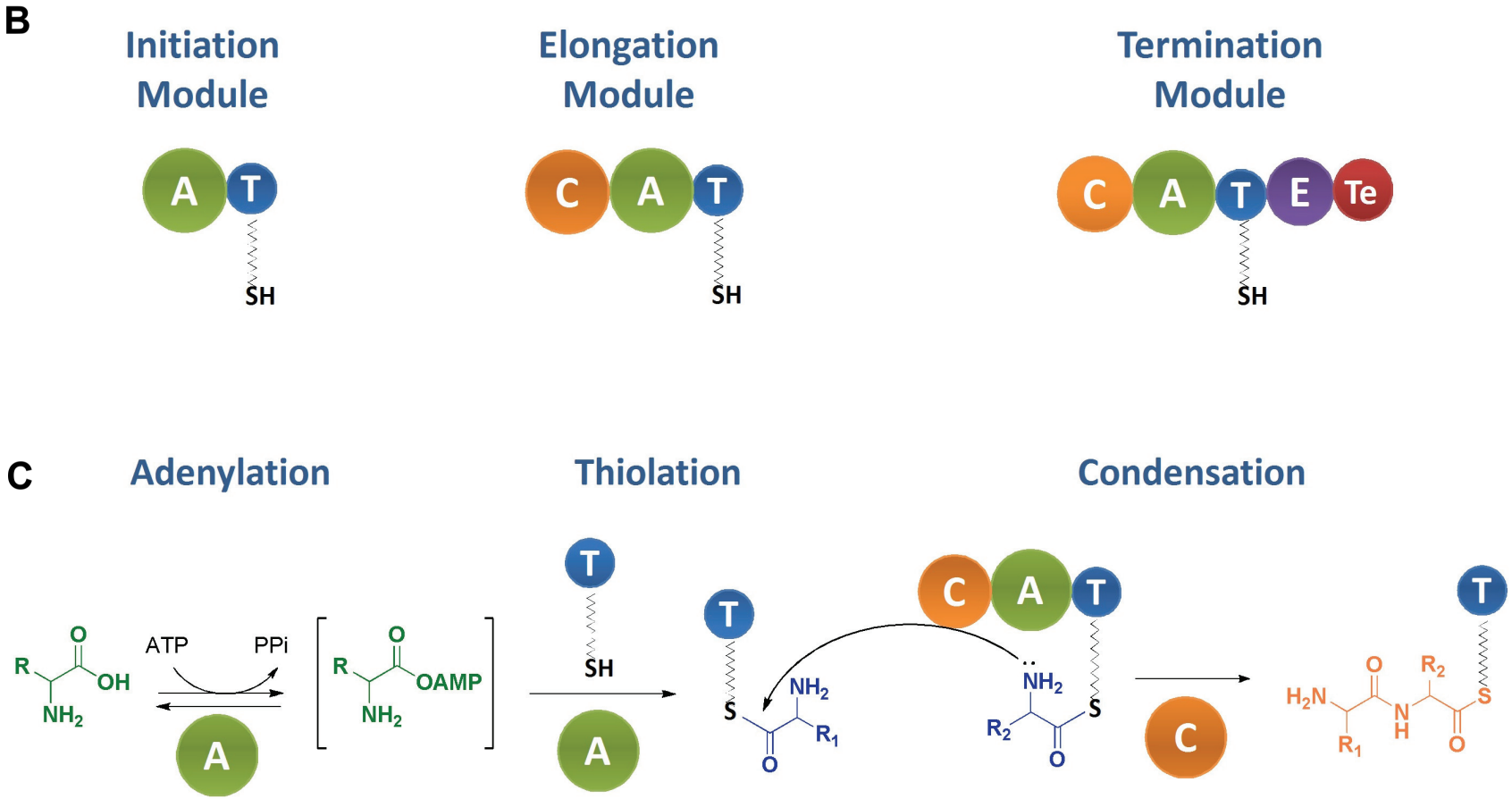


Figure 1.5 (continued) General organization of NRPS assembly line showing (A) conversion of an apo T to a holo T domain by posttranslational phosphopantetheinylation, (B) three modules with multiple domains, and (C) reactions taking place in core domains.

1.4 – SAM-dependent Methyltransferases

Methylation is one of the most common reactions in biology, required in the biosynthesis of hormones and neurotransmitters, DNA maturation, and esterification of carboxylate groups in certain proteins.[19] The reaction is known to occur in two steps of complexation between a donor and an acceptor molecule followed by methyl transfer in an associative mechanism via a single transition state (Figure 1.6*A*).[20] Among known methyl donors, *S*-adenosyl-L-methionine (SAM or AdoMet) is the most prevalently used cofactor in Nature, with a chemical structure shown in Figure 1.6*B*. Upon transfer of the methyl group, AdoMet is transformed to the more stable *S*-adenosyl-L-homocysteine (SAH or AdoHcy).

SAM-dependent methyltransferases catalyze nucleophilic attack of oxygen, nitrogen, sulfur, and carbon nucleophiles on the sp^3 carbon (of the methyl group) in a classic S_N2 fashion, often requiring Mg^{++} as cofactor.[21,22] Structurally, mostly known SAM-dependent methyltransferases with reported crystal structures share similar core structures, consisting of a typical α/β fold similar to the nucleotide-binding motif (Rossman fold).[23] The active site is found between the SAM-binding domain and substrate-binding domain of the enzyme. Most SAM-dependent *O*-methyltransferases catalyze methylation of carboxylate or hydroxyl groups, and *O*-methylation on phosphonic acid has not been reported in the past. Conversion of phosphonates to a methyl ester reduces the negative charge on the molecule making it more efficient for cell diffusion.

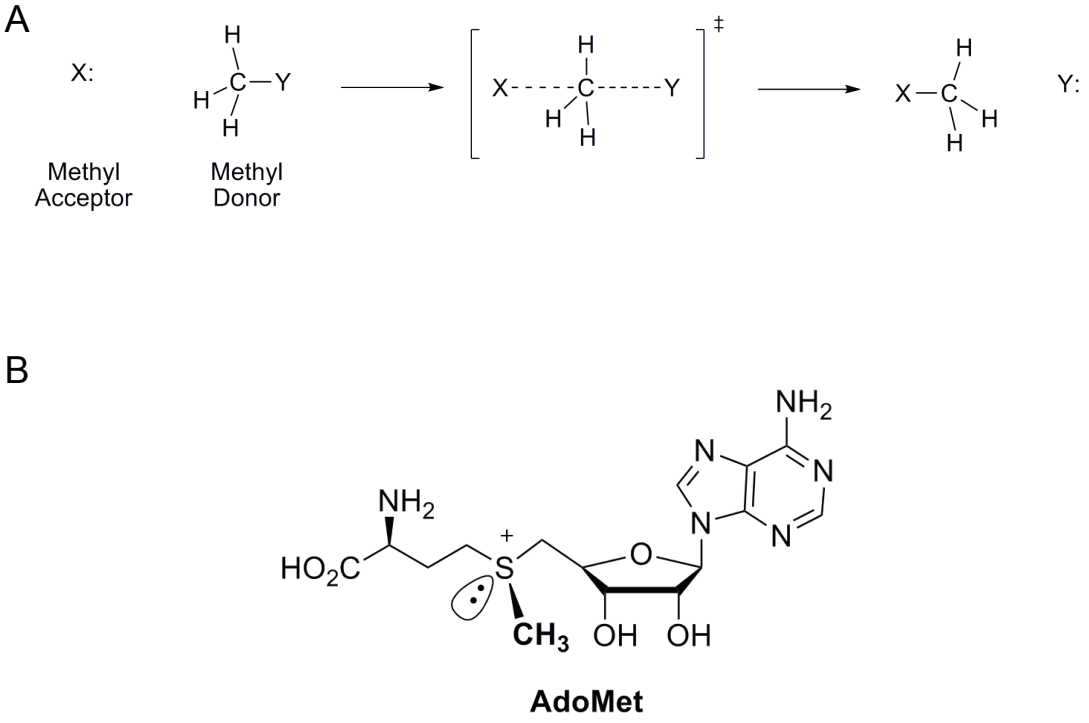


Figure 1.6 Methyl transfer reaction. (A) Concerted mechanism of a methyl transfer reaction (B) The chemical structure of SAM. The stereochemistry at the sulfur atom is important for the biological activity of SAM.[24]

1.5 – Summary

There is a continued need for development of new antibiotics due to ever-increasing antibiotic resistance towards currently available antibiotics. In the past, naturally occurring biologically active compounds have proven to be great drug candidates and still continue to be a major reservoir for novel antibiotics. Recent progress in investigation of the biosynthesis of phosphonate antibiotics has revealed aspects of phosphonate metabolism in microorganisms, and the enzymes involved have provided a wealth of interesting and unique chemistry. Often the peptide bond formation in the phosphonate and phosphinate biosynthesis involves catalysis by NRPSs introducing a vast number of unique nonproteinogenic amino acids. This thesis will discuss the experiments conducted with one such enzyme involved in the biosynthesis of bialaphos as described in Chapter 2. In addition, an interesting SAM-dependent methyl transfer reaction is found in the biosynthesis of the phosphonate dehydrophos possibly to alleviate unfavorable high negative charge in the phosphonic acid. Investigations into the biosynthetic pathway of dehydrophos are presented in Chapter 3 whereas studies of the methyl transferase are described in Chapter 4.

1.6 – References

1. Walsh, C.: *Antibiotics: actions, origins, resistance*. Washington, DC: ASM Press; **2003**.

2. Wright, G. D.: "The antibiotic resistome: the nexus of chemical and genetic diversity." *Nat. Rev. Microbiol.* **2007**, *5*, 175-186.

3. Palumbi, S. R.: "Evolution - Humans as the world's greatest evolutionary force." *Science* **2001**, *293*, 1786-1790.

4. von Nussbaum, F.; Brands, M.; Hinzen, B.; Weigand, S.; Habich, D.: "Antibacterial natural products in medicinal chemistry - Exodus or revival?" *Angew. Chem. Int. Ed.* **2006**, *45*, 5072-5129.

5. Clardy, J.; Fischbach, M. A.; Walsh, C. T.: "New antibiotics from bacterial natural products." *Nat. Biotechnol.* **2006**, *24*, 1541-1550.

6. Projan, S. J.: "Why is big Pharma getting out of antibacterial drug discovery?" *Curr. Opin. Microbiol.* **2003**, *6*, 427-430.

7. Sabol, K.; Patterson, J. E.; Lewis, J. S.; Owens, A.; Cadena, J.; Jorgensen, J. H.: "Emergence of daptomycin resistance in *Enterococcus faecium* during daptomycin therapy." *Antimicrob. Agents Chemother.* **2005**, *49*, 1664-1665.

8. Metcalf, W. W.; van der Donk, W. A.: "Biosynthesis of phosphonic and phosphinic acid natural products." *Annu. Rev. Biochem.* **2009**, *78*, 65-94.

9. Borisova, S. A.; Circello, B. T.; Zhang, J. K.; van der Donk, W. A.; Metcalf, W. W.: "Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by *Bacillus subtilis* ATCC6633." *Chem. Biol.* **2010**, *17*, 28-37.

10. Woodyer, R. D.; Shao, Z. Y.; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A. V.; Metcalf, W. W.; Van der Donk, W. A.; Zhao, H. M.: "Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster." *Chem. Biol.* **2006**, *13*, 1171-1182.

11. Blodgett, J. A. V.; Thomas, P. M.; Li, G. Y.; Velasquez, J. E.; van der Donk, W. A.; Kelleher, N. L.; Metcalf, W. W.: "Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide." *Nat. Chem. Biol.* **2007**, *3*, 480-485.

12. Eliot, A. C.; Griffin, B. M.; Thomas, P. M.; Johannes, T. W.; Kelleher, N. L.; Zhao, H. M.; Metcalf, W. W.: "Cloning, expression, and biochemical characterization of *Streptomyces rubellomurinus* genes required for biosynthesis of antimalarial compound FR900098." *Chem. Biol.* **2008**, *15*, 765-770.

13. Johannes, T. W.; DeSieno, M. A.; Griffin, B. M.; Thomas, P. M.; Kelleher, N. L.; Metcalf, W. W.; Zhao, H. M.: "Deciphering the late biosynthetic steps of antimalarial compound FR-900098." *Chem. Biol.* **2010**, *17*, 57-64.

14. Ntai, I.; Bachmann, B. O.: "Identification of ACE pharmacophore in the phosphonopeptide metabolite K-26." *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3068-3071.

15. Blodgett, J. A. V.; Zhang, J. K.; Metcalf, W. W.: "Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736." *Antimicrob. Agents Chemother.* **2005**, *49*, 230-240.

16. Circello, B. T.; Eliot, A. C.; Lee, J.-H.; van der Donk, W. A.; Metcalf, W. W.: "Molecular cloning and heterologous expression of the dehydrophos biosynthetic gene cluster." *Chem. Biol.* **2010**, *17*, 402-411.

17. Marahiel, M. A.; Stachelhaus, T.; Mootz, H. D.: "Modular peptide synthetases involved in nonribosomal peptide synthesis." *Chem. Rev.* **1997**, *97*, 2651-2673.

18. Grunewald, J.; Marahiel, M. A.: "Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides." *Microbiol Mol Biol R* **2006**, *70*, 121-+.

19. Cantoni, G. L.: "Biological methylation: Selected aspects." *Annu. Rev. Biochem.* **1975**, *44*, 435-451.

20. Wang, H. B.; Hase, W. L.: "Kinetics of F-+CH_3Cl S_N2 nucleophilic substitution." *J. Am. Chem. Soc.* **1997**, *119*, 3093-3102.

21. Coward, J. K.; Slisz, E. P.; Wu, F. Y. H.: "Kinetic studies on catechol *O*-methyltransferase: Product inhibition and nature of catechol binding-site." *Biochemistry* **1973**, *12*, 2291-2297.

22. Hegazi, M. F.; Borchardt, R. T.; Schowen, R. L.: "S_N2-like transition-state for methyl transfer catalyzed by catechol-*O*-methyltransferase." *J. Am. Chem. Soc.* **1976**, *98*, 3048-3049.

23. Vidgren, J.; Svensson, L. A.; Liljas, A.: "Crystal structure of catechol *O*-methyltransferase." *Nature* **1994**, *368*, 354-358.

24. Iwig, D. F.; Booker, S. J.: "Insight into the polar reactivity of the onium chalcogen analogues of *S*-adenosyl-L-methionine." *Biochemistry* **2004**, *43*, 13496-13509.

CHAPTER 2: STUDY OF A NONRIBOSOMAL PEPTIDE SYNTHETASE INVOLVED IN PHOSPHINOTHRICIN TRIPEPTIDE BIOSYNTHESIS[*]

2.1 – Introduction

Phosphinothricin tripeptide (PTT, also called bialaphos, Figure 2.1) is produced by *Streptomyces viridochromogenes* and *Streptomyces hygroscopicus*.[21] This natural product consists of an unusual phosphinic acid phosphinothricin (PT), which contains a unique carbon-phosphorus-carbon moiety, and two alanines. PTT enters a bacterial cell in a Trojan-horse manner via an oligopeptide transporter. Once inside the cell, the active PT is released by universal peptidase activity.[22]

This unusual *P*-methylated amino acid is a structural mimic of a tetrahedral intermediate formed during nucleophilic addition of ammonia to glutamic acid in the glutamine synthetase reaction.[23] Complete inhibition of glutamine synthetase activity results in ammonia build-up that is toxic to plants, and in bacteria this inhibition leads to glutamine shortage (Figure 2.2). PTT is widely used as an herbicide under the tradename Herbiace and the ammonium salt of PT is the active ingredient in BASTA. Transgenic plants engineered to contain the PT-resistance gene (*pat* in *S. viridochromogenes*,[24] Figure 2.1, or *bar* in *S. hygroscopicus*[25]) have found widespread use in agriculture in the form of LibertyLink genetically modified crops such as corn, cotton, soybean and canola.

[*] Reproduced in part with permission from "In Vitro Characterization of a Heterologously Expressed Nonribosomal Peptide Synthetase Involved in Phosphinothricin Tripeptide Biosynthesis." *Biochemistry*, **2009**, *48*, 5054-5056. Copyright 2009 by the American Chemical Society.

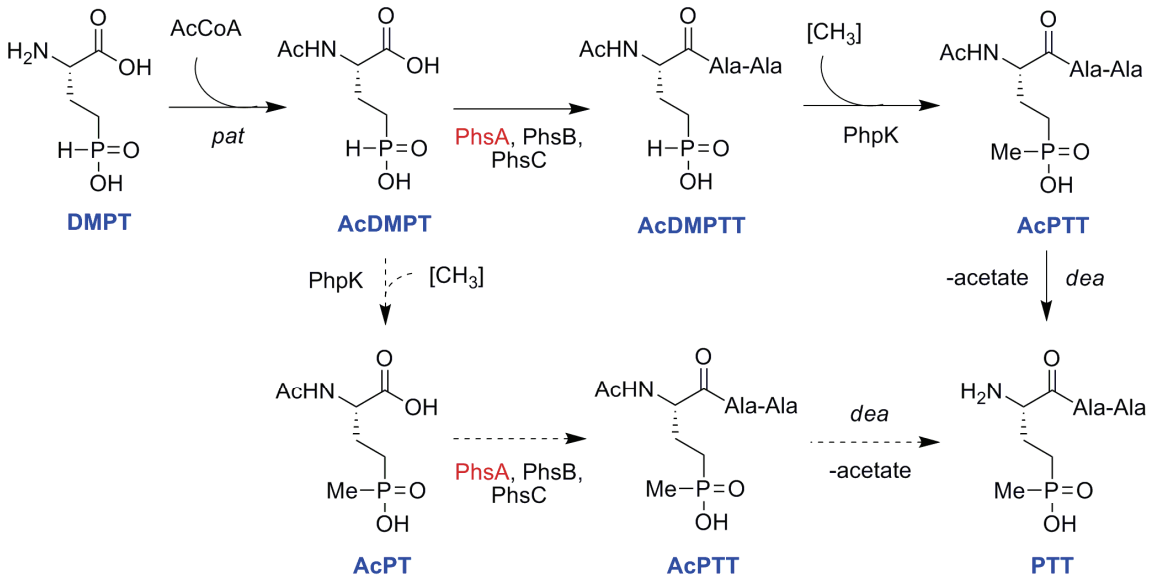


Figure 2.1 The currently accepted biosynthetic pathway for PTT is shown with solid arrows. An alternative pathway in which AcDMPT is first methylated and then incorporated into the tripeptide is depicted in dashed arrows.

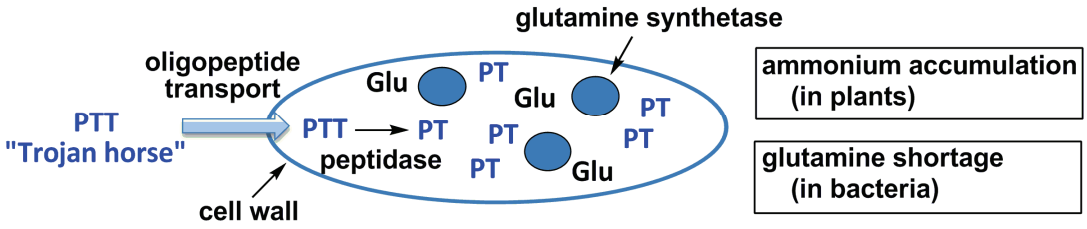


Figure 2.2 Schematic representation of PTT mode of action in both plant and bacterial cells.

In recent years, the PTT biosynthetic gene cluster in *S. viridochromogenes* has been fully sequenced[26,27] and part of the cluster of *S. hygroscopicus* has also been reported.[28,29] The final steps in PTT biosynthesis include tripeptide assembly as well as installation of the methyl group on the phosphorus atom (Figure 2.1).[30-32] PhpK-catalyzed *P*-methylation, which endows PTT its characteristic phosphinate moiety, has been of great interest due to the structural importance and agricultural applications of this compound.[33] However, the exact timing of the methyl transfer step, whether before, during or after peptide assembly, is currently unknown.[34] A mutant in *phsA,* encoding one of three nonribosomal peptide synthetase proteins in the PTT gene cluster (PhsA, PhsB, and PhsC),[30,32,35] accumulated *N*-acetyldemethylphosphinothricin (AcDMPT) suggesting that this phosphinate amino acid is the substrate for PhsA.[36,37]

From the amino acid sequence, it has been suggested that PhsA is an initiation module in the nonribosomally synthesized tripeptide.[30] From the sequence homology analysis, PhsA was shown to be comprised of an adenylation (A) domain that activates the *N*-terminal phosphinic acid scaffold to its corresponding aminoacyl adenylate, a thiolation (T) domain that covalently tethers the amino acid to the phosphopantetheine (PPant) arm of the peptide carrier protein, and a putative *N*-terminal thioesterase (Te) sequence motif of unknown function (Figure 2.3*A,C*).[35] After charging of the phosphinate amino acid onto PhsA, the two peptide bond forming events are catalyzed by the condensation domains of PhsB and PhsC, each loaded with alanine (Figure 2.3*B*).[30,32] The order of peptide bond formation and how the final product is released from the peptide carrier protein (PCP) of PhsB/PhsC is currently not clear. The final release step could involve the *N*-terminal thioesterase (Te) motif of PhsA or two stand-alone type II thioesterases in the PTT gene cluster (Figure 2.3*C*).[35]

Keller and Wohlleben and their coworkers isolated the native PhsA enzyme from *S. viridochromogenes* in small quantities.[30] Sufficient protein was obtained to study the activity of the enzyme by ATP-pyrophosphate (PP_i) exchange assay, demonstrating that AcDMPT was utilized most efficiently under the conditions employed, but that *N*-acetyl phosphinothricin (AcPT) was also a substrate. Attempts to express the enzyme in *Escherichia coli* to obtain larger quantities of protein were not successful. Collectively, these studies suggest the order of events shown in solid arrows in Figure 2.1. The observed activity of charging PhsA with AcDMPT suggests the possibility that the methylation step is a tailoring reaction that takes place while AcDMPT is tethered to the PPant arm of PhsA (or PhsB/PhsC).

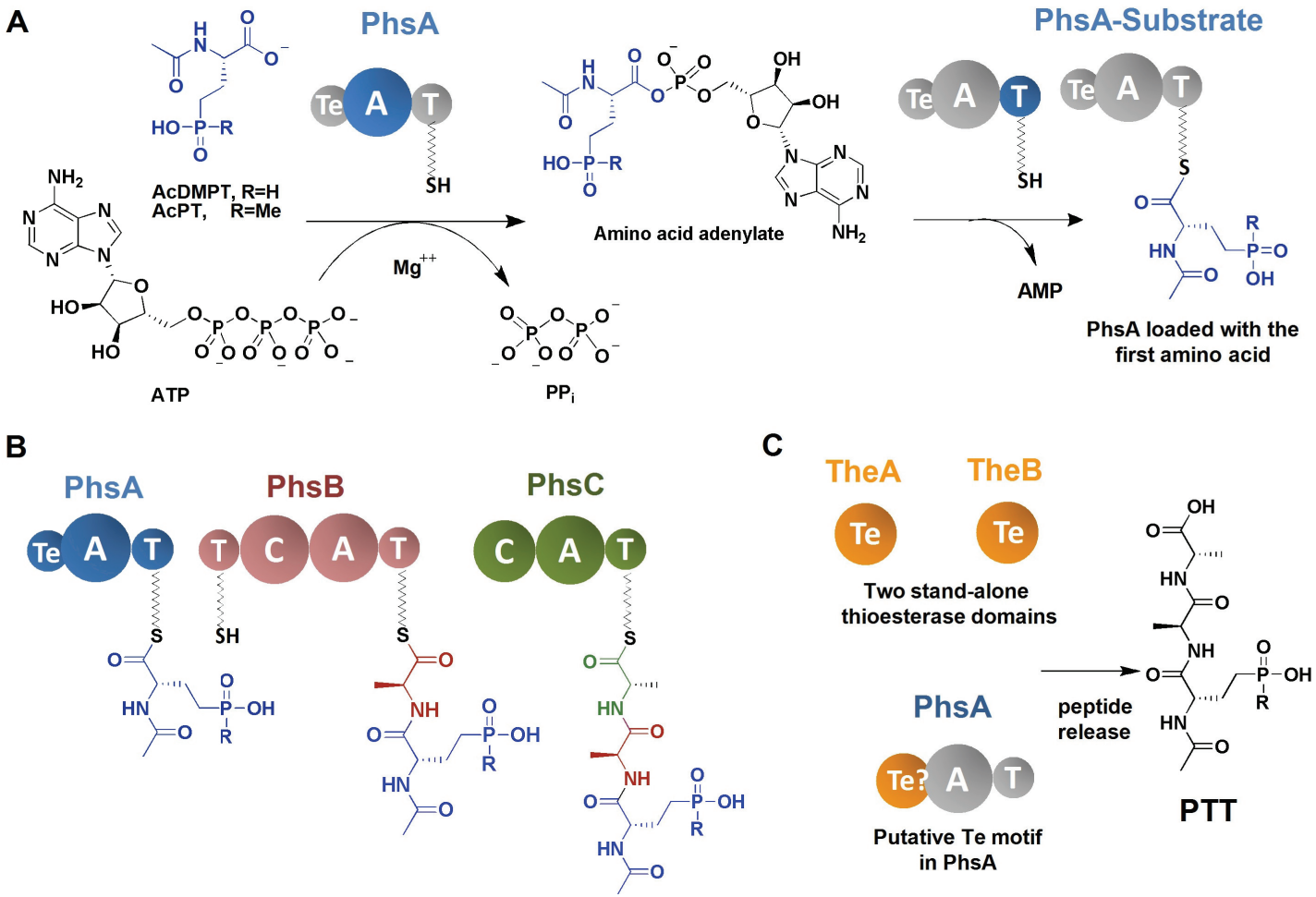


Figure 2.3 Overview of PTT amide bond formation. (A) PhsA, the initiation module contains an adenylation (A) and a thiolation (T) domain, and is responsible for selection, activation, and loading of the phosphinic amino acid. (B) PhsB and PhsC are responsible for incorporating two alanines into PTT. (C) Two stand-alone Te domains, TheA and/or TheB or a putative Te motif in the *N*-terminus of PhsA may be involved in PTT release.

In order to investigate the molecular details for this step, the possibility to express active PhsA in *E. coli* has been revisited. In this study, we show that a fusion protein of PhsA to maltose binding protein (MBP) was expressed in soluble form and that it was active in ATP-PP_i exchange assays. *In vitro* phosphopantetheinylation of apo-PhsA by PPant transferase, Sfp, from *Bacillus subtilis* was investigated using various detection methods. The kinetic parameters for ATP-PP_i exchange were determined for a series of substrates including AcDMPT for which synthetic preparation had not been reported before the start of this work. In addition, formation of a covalent thioester linkage of loaded amino acid substrate to the PPant group of the T domain was assessed using Fourier-Transform mass spectrometry (FTMS). The results of this study as well as previous studies by Seto and coworkers[33,34] suggest that an alternative pathway in which AcDMPT is first methylated cannot be ruled out (dashed arrows, Figure 2.1).

2.2 – Results

2.2.1. Heterologous Expression of PhsA

Initial efforts to express PhsA as an *N*-terminal His_6-tagged protein in *E. coli* BL21 (DE3) by cloning the *phsA* gene from genomic DNA of *S. viridochromogenes* DSM 40736 produced insoluble protein. Optimization of expression conditions in an attempt to produce more soluble PhsA included variations in induction conditions as well as co-expression of PhsA with chaperone proteins using commercially available pG-KJE8 plasmid, which co-expresses the DnaK-DnaJ-GrpE-GroES-GroEL chaperone proteins.[38] Other attempts in collaboration with the Kelleher group including conjugative transformation of *S. lividans* with the plasmid and cloning into a *C*-terminal His_6-tag

system met with little success. It is interesting to note, however, that His_6-PhsA with an undesired mutation at the active site serine residue (S579F) obtained from one of the clones turned out to be more soluble and its purification could be achieved using Ni-affinity column chromatography (data not shown). It is tempting to speculate that the mutation results in a fold that resembles the holo form of the enzyme, which is expected to be more stable. Finally, the *phsA* gene was inserted into the pMAL-c2x vector to introduce MBP as a fusion at the *N*-terminus. Overexpression of PhsA as an MBP fusion in *E. coli* BL21 resulted in the production of a significant amount of soluble protein. However, after purification of the MBP-fusion protein by amylose affinity chromatography, the MBP-PhsA was isolated in a partially pure form. Truncated proteins were visualized via sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), which were initially suspected to originate from premature termination of *phsA* translation due to rare codons in *phsA* (namely CCC for proline and CGG for arginine) that *E. coli* does not translate well. Therefore *E. coli* Rosetta 2 BL21 strain, which supplements tRNAs for these rare codons, was used for expression instead. Nevertheless, the truncated proteins still remained before and after the purification. Their continued presence suggests that they may have formed from the proteolytic cleavage of PhsA. To improve the purification, the *malE-phsA* fragment was cloned into pET28b and MBP-PhsA-His_6 protein was heterologously expressed in *E. coli*. Rapid purification of the resulted fusion protein by Ni-NTA chromatography followed by HiTrap Q HP ion-exchange chromatography yielded 7 mg of intact protein per liter of culture (Figure 2.4).

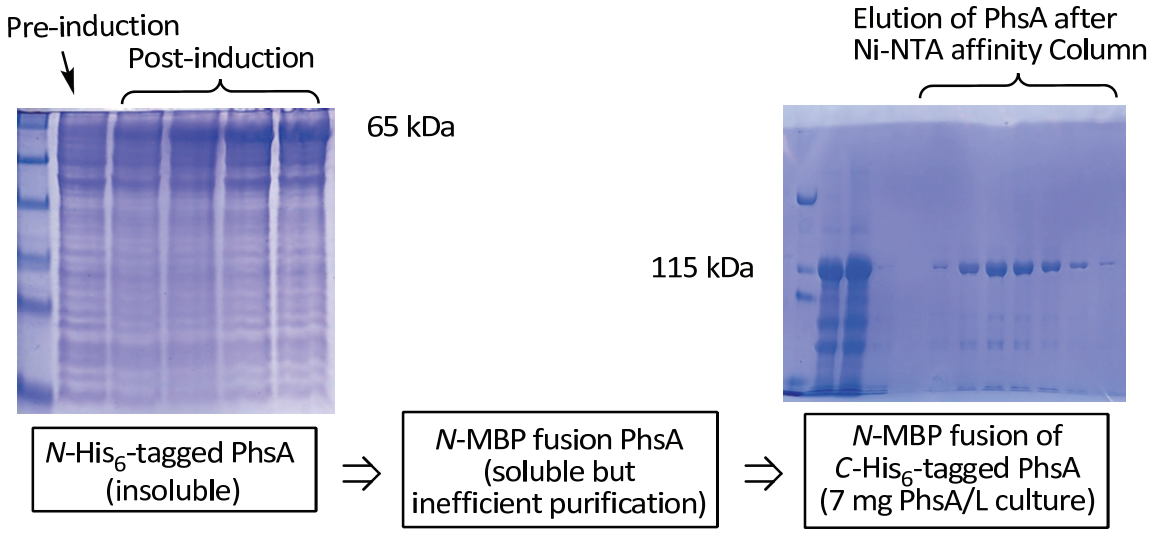


Figure 2.4 Heterologous expression of PhsA in a soluble form. A construct which expresses an *N*-terminal MBP fusion of *C*-terminally His_6-tagged PhsA allowed isolation of PhsA in a soluble form.

2.2.2. In Vitro Phosphopantetheinylation of PhsA

As isolated, the recombinantly expressed MBP-PhsA-His_6 protein did not contain detectable amounts of the phosphopantetheine arm as monitored by FTMS. Two experiments were performed to examine the *in vitro* phosphopantetheinylation of MBP-PhsA-His_6 by Sfp (Figure 2.5*A*), a highly promiscuous PPant transferase from the surfactin biosynthetic operon of *B. subtilis*. In the first experiment[†], the fluorophore-carrying coenzymeA (CoA) analogue BODIPY-FL-*N*-(2-aminoethyl)maleimidyl-*S*-CoA (BODIPY-CoA, Figure 2.5*B*) was employed to install a fluorescently labeled PPant group on the active site serine residue in the PhsA T domain, which could be readily detected spectrofluorometrically. The samples of PhsA that were treated with Sfp and BODIPY-CoA were analyzed by SDS-PAGE and visualized by coomassie blue staining, as well as by fluorescence spectroscopy (Figure 2.5*C*). When Sfp was co-incubated with PhsA in the presence of BODIPY-CoA, PhsA could be spectrofluorophotometrically detected, showing the successful transfer of the PPant group onto PhsA. In a second experiment ^{3}H-acetyl-CoA was utilized as co-substrate, resulting in a time-dependent increase in the level of radiolabeled phosphopantetheinyl group from ^{3}H-acetyl CoA to MBP-PhsA-His_6. Maximum incorporation of tritium was achieved after less than 15 min and was nearly stoichiometric (Figure 2.5*D*), confirming that Sfp converts apo-MBP-PhsA-His_6 to its holo form with high efficiency.

[†] *In vitro* phosphopantetheinylation experiment by in-gel fluorescence labeling was performed in collaboration with Bradley S. Evans from the laboratory of Professor Neil L. Kelleher (Department of Chemistry, UIUC).

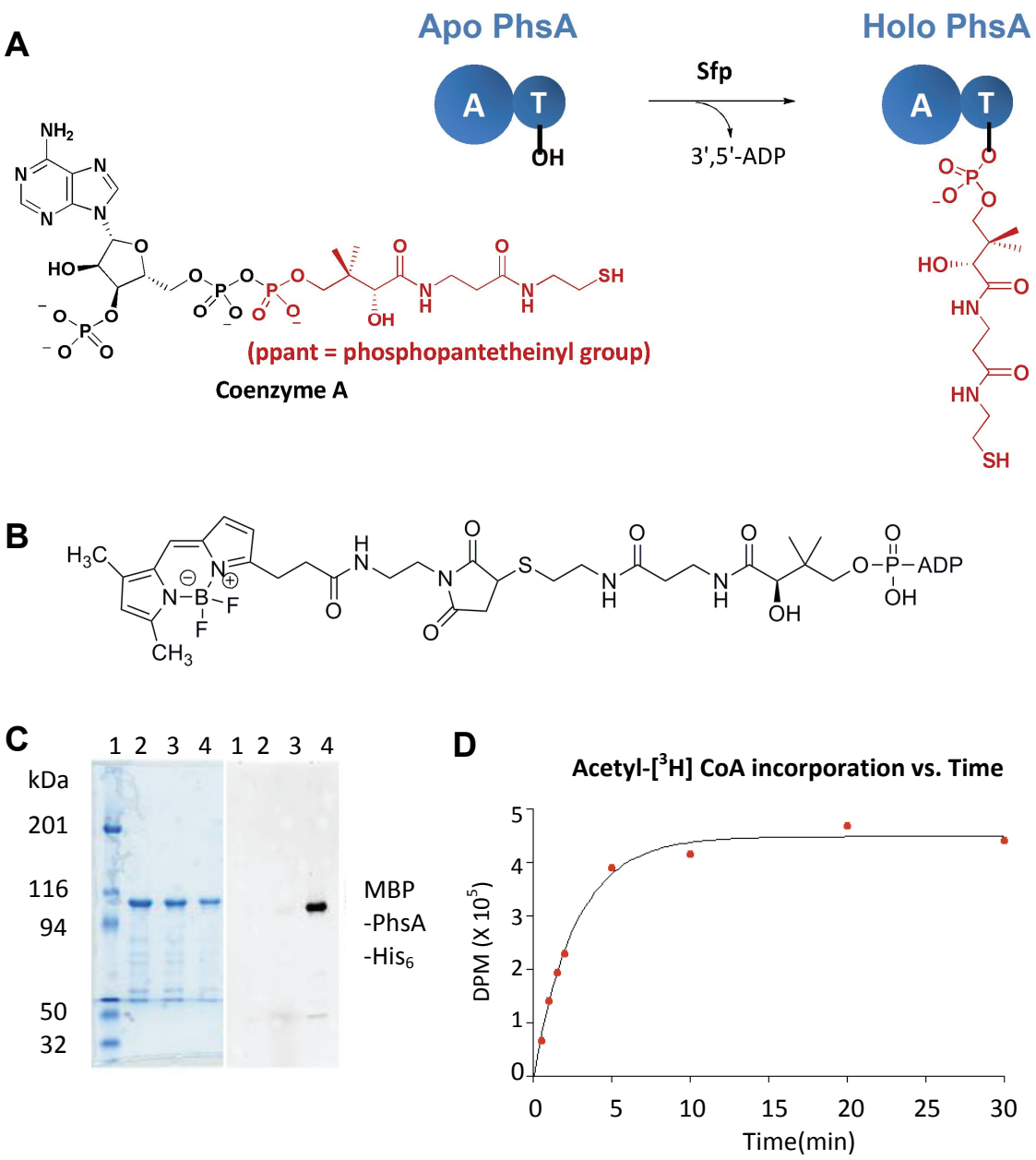


Figure 2.5 *In vitro* phosphopantetheinylation of MBP-PhsA-His_6. (A) Conversion of apo PhsA to its holo form by the PPTase Sfp. **(**B) Chemical structure of BODIPY-CoA. (C) In-gel fluorescent labeling. SDS-PAGE of reaction mixture after coomassie blue staining (left) and when visualized at 520 nm (right). Lane 1, protein standard; lane 2, negative control with only MBP-PhsA-His_6; lane 3, MBP-PhsA-His_6 treated with BODIPY-CoA; lane 4, MBP-PhsA-His_6 treated with BODIPY-CoA and Sfp. (D) Time-dependent incorporation of tritium into MBP-PhsA-His_6 during incubation with Sfp.

2.2.3. Syntheses of PhsA Substrate Analogues

The previous investigation of the substrate scope of native purified PhsA showed that both AcDMPT and AcPT exhibited significant ATP-pyrophosphate exchange activities with AcPT having slightly lower activity (74%).[30] These studies were performed at a substrate concentration of 667 μM (U. Keller, personal communication). In order to better understand the substrate scope of this enzyme at physiological conditions, it was desired to examine the initial rates of the enzymatic reaction at varying concentrations of the substrate analogs.

In this study, we obtained sufficient amounts of protein for full kinetic characterization with potential substrates and substrate analogues (Figure 2.6). These compounds were prepared by chemical synthesis by Dr. Gongyong Li[‡] from the van der Donk group and myself. AcDMPT was synthesized in seven steps starting from L-methionine in 11% overall yield (Figure 2.7*A*). In this synthesis, L-methionine was first alkylated with bromoacetic acid followed by attack of the carboxylate oxygen onto the *β*-carbon of the sulfonium salt, resulting in homoserine lactone **1**. This lactone was brominated with hydrobromic acid in the presence of acetic acid to produce bromide **2**. The carboxyl group of this compound was converted to its methyl ester **3** in hydrochloric methanol, followed by acetylation of the amino group to yield fully protected compound **4**. The bromide was treated with sodium iodide, and the resulting iodide **10** was reacted with bis(trimethylsilyl)phosphonite (BTSP) generated *in situ* from the reaction of ammonium hypophosphite and hexamethyldisilazane.[39] The ammonium salt **11** was reacted with sodium hydroxide followed by purification on Dowex resin to obtain the

[‡] Dr. Gongyong Li prepared AcDMPT. I participated in the synthesis of AcDMPT by preparing the intermediate **3** for the downstream steps performed by Dr. Li.

desired compound **12**. Due to the low reaction yield, difficulty in purification of the product, as well as involvement of the highly pyrophoric compound BTSP in the phosphination step by Arbuzov-type reaction, an alternative method for AcDMPT preparation was explored (Figure 2.7*B*). In this method, radical addition of phosphite to the alkene **7** catalyzed by triethylborane, was achieved.[40] This step proceeded with a high yield of 83% and included a more efficient purification step, and thus for large scale production of AcDMPT required in the enzymatic assay, the second route (Figure 2.7*B*) was employed.

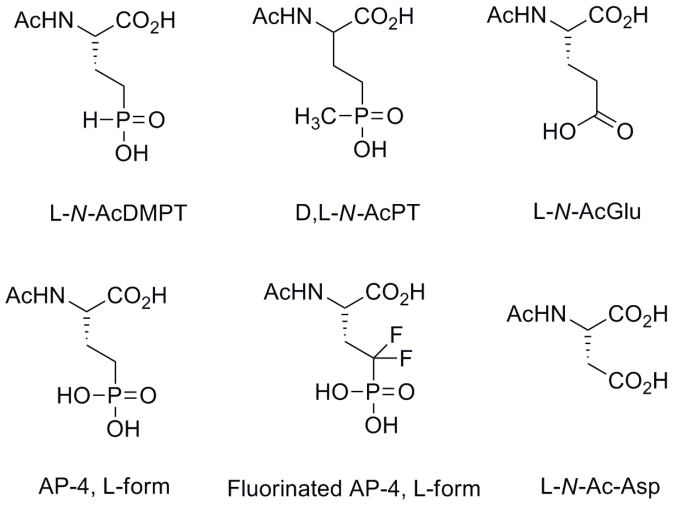


Figure 2.6 Acetylated amino acid substrate analogues used in the PhsA enzymatic assay.

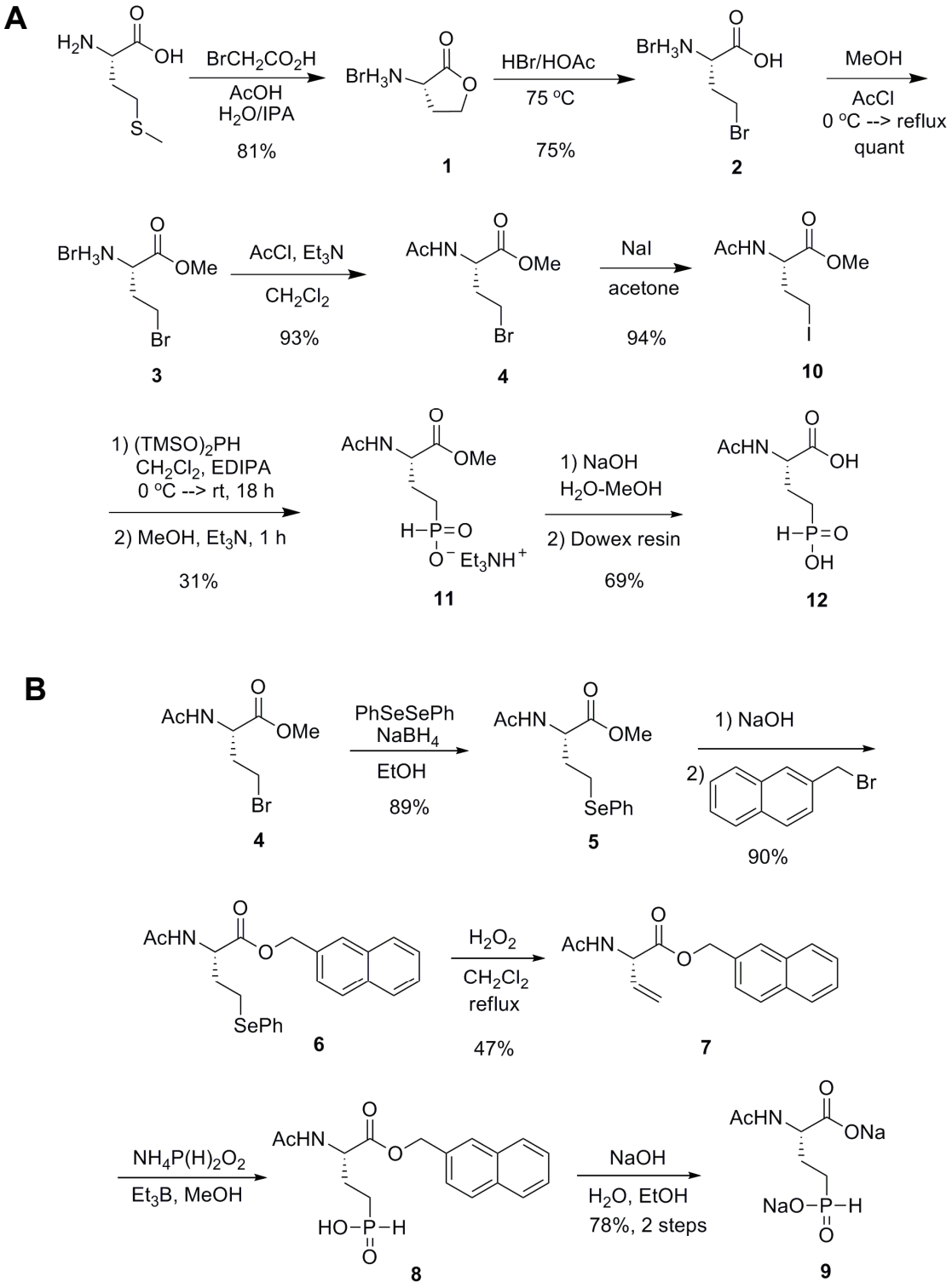


Figure 2.7 Synthesis of AcDMPT. (A) Original reaction scheme in which Arbuzov-type reaction of the iodide with BTSP was used. (B) Alternative method in which radical addition of phosphite to the alkene **7** was employed to introduce the phosphinate.

AcPT was prepared by acetylation of commercially available D,L-PT by reaction with acetic anhydride and potassium carbonate. Previous studies have shown that the acetylated form of the amino acid substrate is essential for PhsA activity,[30,33] and thus various other amino acids which are structurally similar to AcDMPT, including L-(+)-2-amino-4-phosphonobutyric acid (L-AP-4), L-glutamate, L-aspartate, and fluorinated L-AP-4[§], were all acetylated (Figure 2.6) prior to testing for activity with PhsA. With the exception of AcPT, which was prepared as a racemic mixture, all substrate analogues were enantiomerically pure.

2.2.4. Determination of the Substrate Scope of Adenylation Activity by PhsA

The substrate set was used to probe PhsA activity by the ATP-PP_i exchange assay using a published protocol.[41] At 2 mM concentrations, L-*N*-AcAsp, L-*N*-AcAP4 and fluorinated L-*N*-AcAP4 showed very low activity and were not investigated further. On the other hand, L-*N*-AcGlu, L-AcDMPT, and D,L-AcPT showed significant ATP-PP_i exchange activity and the substrate concentration dependence was further investigated. The initial exchange rates (<13% conversion) were plotted against substrate concentration as depicted in Figure 2.8. It has been noted for other investigations into the kinetics of the ATP-PP_i exchange reaction that the unlabeled reactants very rapidly establish an equilibrium and hence that the observed rate of exchange is the equilibrium rate of production of radiolabeled ATP from labeled PP_i (Figure 2.9). As such, although the ATP-PP_i exchange reaction provides insights into enzyme specificity, it is not straightforward to interpret its kinetics in the framework of the Michaelis-Menten equation.[42] However, because of the difficulties associated with measuring direct rates of

[§] Dr. Xingang Zhang from the van der Donk group provided fluorinated L-AP-4.

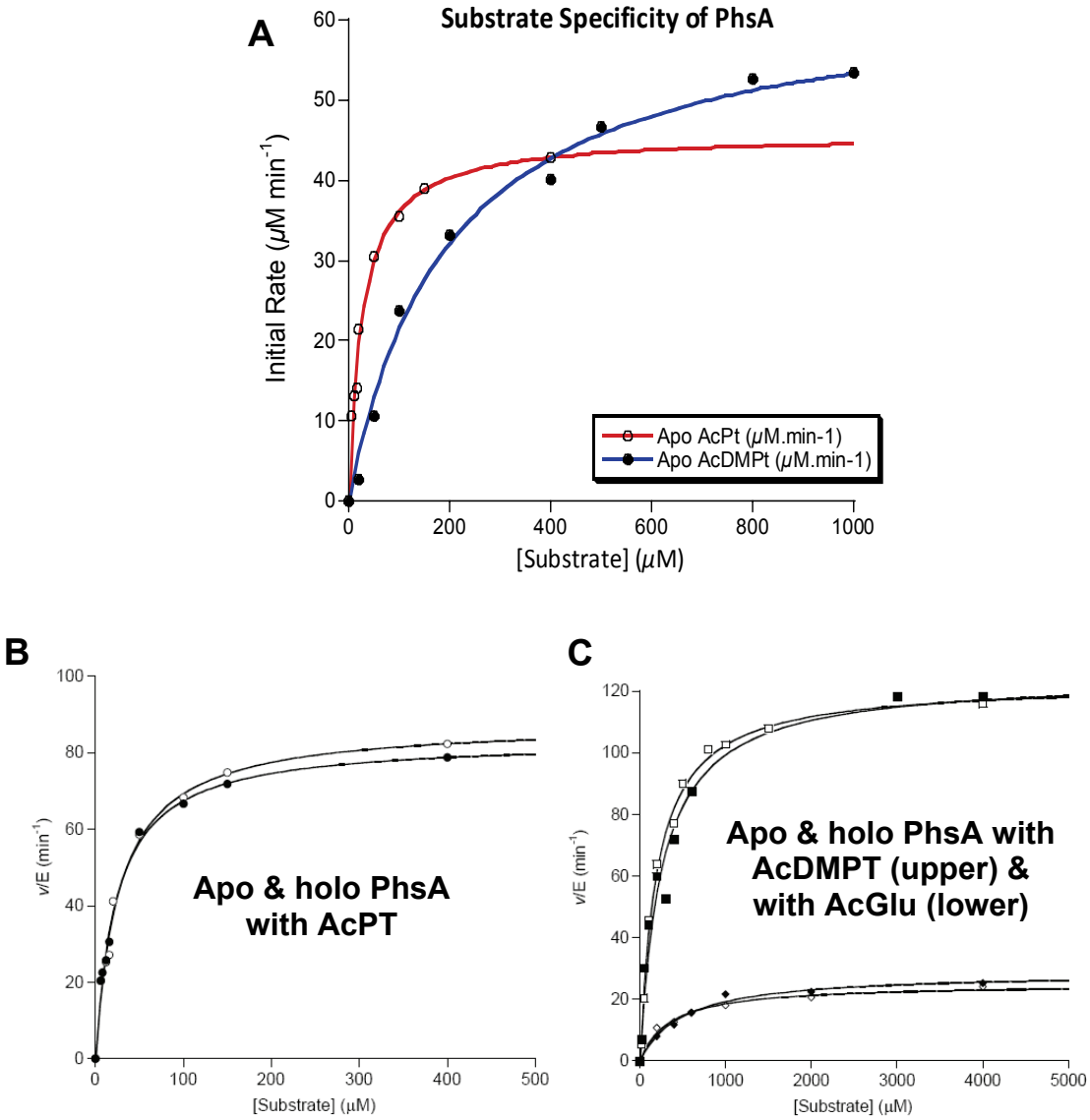


Figure 2.8 Concentration dependence of the observed initial rates of amino acid-stimulated ATP-pyrophosphate exchange. The lines were obtained by fitting the data to the Michaelis-Menten equation. (A) Initial rate of D,L-AcPT (red curve) compared to that of L-AcDMPT (blue curve) when the reaction was carried out by apo-MBP-PhsA-His_6. When the substrate concentration was less than 400 μM, the initial rate of D,L-AcPT is higher than that of L-AcDMPT. (B) Apo-MBP-PhsA-His_6 (○) and holo-MBP-PhsA-His_6 (●) with D,L-AcPT. (C) Apo-MBP-PhsA-His_6 with L-AcDMPT (□) or L-*N*-AcGlu (◇), and holo-MBP-PhsA-His_6 with L-AcDMPT (■) or L-*N*-AcGlu (◆).

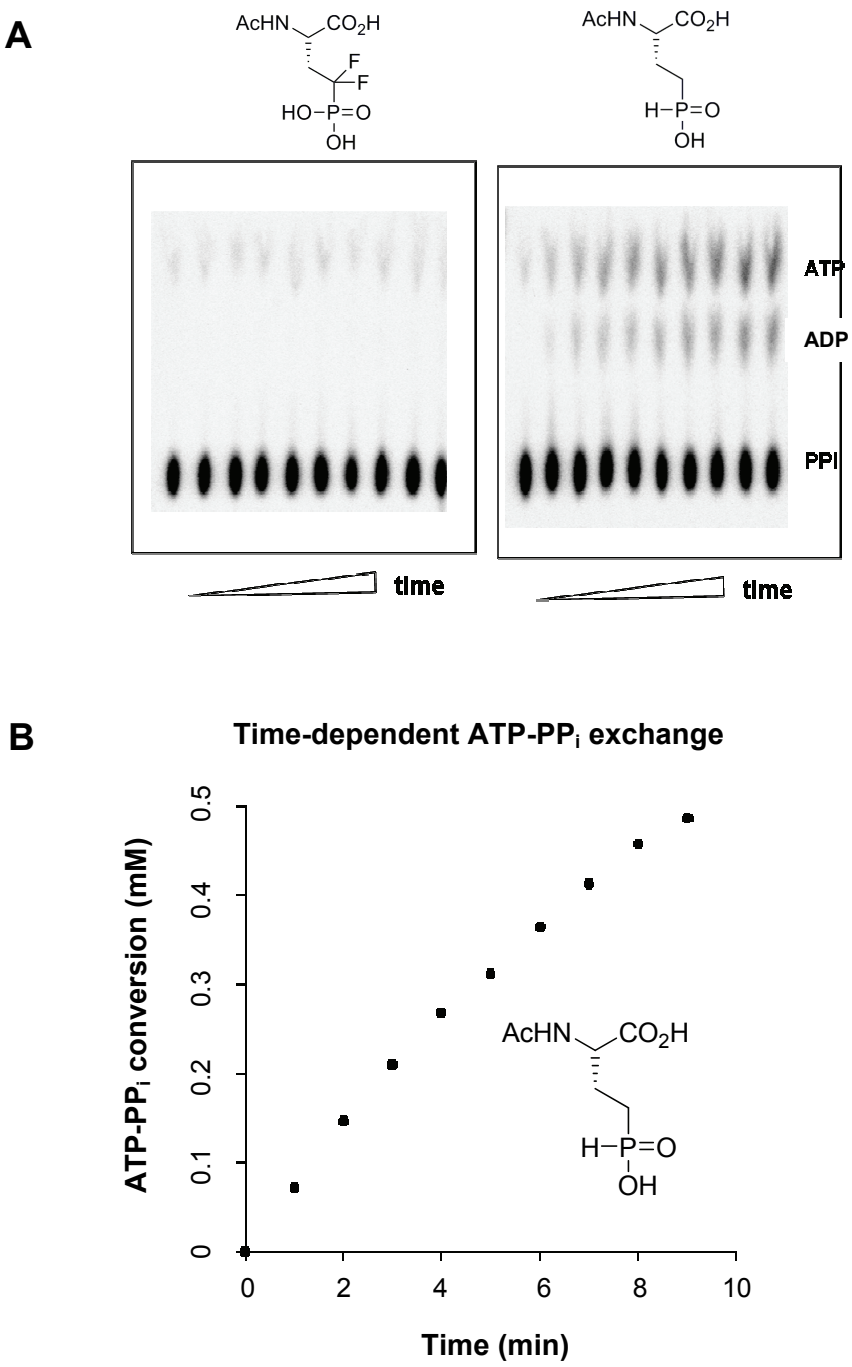


Figure 2.9 The ATP-pyrophosphate exchange assay. (A) Formation of the radiolabeled ATP monitored by separation of the radioactive species (i.e. ATP, ADP, and PP_i) by thin layer chromatography (TLC) followed by analysis with a phosphorimager. When a poor substrate analogue (fluorinate L-AP-4, on the left) was used little amount of radiolabeled ATP was detected. When AcDMPT was used instead (on the right), formation of radiolabeled ATP was very high. (B) The amount of radiolabeled ATP was plotted against time to obtain the initial rate.

adenylation due to instability of the product as well as very tight binding of the aminoacyl adenylate by A-domains,[43,44] the ATP-PP_i exchange rate has been frequently used to provide insights into the substrate selectivities of adenylation.[43-48] Importantly, in cases where adenylation could be measured directly and were compared with the rates of ATP-PP_i exchange, very similar substrate selectivities were observed.[43]

In addition to the expected exchange of ^{32}P from pyrophosphate into ATP, small amounts of radiolabeled ADP and trace amounts of P_i were also observed, as has been seen in other studies of adenylation domains. ADP is typically not identified using the charcoal adsorption method as both ATP and ADP are adsorbed to the charcoal. The formation of these additional products was dependent on both PhsA and substrate. Inclusion of these radiolabeled products resulted in the same trends (larger apparent k_{cat} for AcDMPT, smaller apparent K_m for AcPT) as omitting them from the analysis.[42] For the system under investigation here, at high concentrations, the observed exchange rate for AcDMPT (125 min^{-1}, Table 2.1) was larger than that for AcPT (84 min^{-1}), consistent with previous results.[30] However, surprisingly the apparent K_m values of the two potential natural substrates, AcPT and AcDMPT, showed that at low substrate concentrations AcPT is the preferred substrate for amino acid-stimulated ATP-PP_i exchange (Table 2.1). This preference is even more pronounced when considering that AcPT is a mixture of enantiomers whereas AcDMPT was stereochemically pure. The ability of the A domain to recognize the substrates and activate them by adenylation was not dependent on the state of the T domain (apo vs. holo, Figure 2.8).

Table 2.1 Kinetic parameters for *N*-acetyl amino acid-stimulated ATP-PP_i exchange reaction by MBP-PhsA-His_6 in its apo and holo forms.

Substrate	Recom-binant PhsA	k_{cat} (min^{-1})	K_m (μM)	k_{cat} / K_m (μM^{-1} min^{-1})	Endo-genous enzyme[30]
(±)-AcPT	apo[a]	87.9 ± 2.9	26.9 ± 2.9	3.27 ± 0.37	
	holo[b]	83.5 ± 2.0	23.7 ± 1.9	3.52 ± 0.29	74%
L-AcDMPT	apo	122.9 ± 2.9	198 ± 18	0.62 ± 0.06	
	holo	124.7 ± 6.5	254 ± 44	0.49 ± 0.09	100%
N-AcGlu	apo	25.0 ± 1.1	349 ± 55	0.0717 ± 0.011	
	holo	28.8 ± 1.5	493 ± 82	0.0584 ± 0.010	11%

[a] The apo form of the enzyme was not treated with Sfp and CoA to install the phosphopantetheine arm.
[b] The holo form has been treated with Sfp and CoA as discussed in the text.

2.2.5. Thiotemplate Analysis of Amino-Acid Loading on PhsA by FTMS

The substrate specificity of holo-MBP-PhsA-His_6 was further analyzed by competition experiments using FTMS in collaboration with Bradley S. Evans from the laboratory of Prof. Neil L. Kelleher (Department of Chemistry, UIUC). The protein (12.4 μM) was incubated with a 1:1 mixture of AcDMPT and AcPT at 25, 250, and 450 μM of each amino acid in the presence of 1 mM ATP. After 1 h, the protein was analyzed by FTMS using the PPant ejection assay.[49,50] The ions corresponding to PPant loaded with AcPT (466.1778) and AcDMPT (452.1616) were observed (Figure 2.10*A*) as well as the parent PPant ejection ion (261.1268). In two independent experiments, the ratios of the

peak intensities of AcPT-PPant and AcDMPT-PPant compounds favor the former at all three substrate concentrations. The presence of the parent ion[51] (261.1267, data not shown here) and its unknown origin (holo PhsA, AcPT−PhsA, or Ac-DMPT−PhsA) complicate interpretation, but the FTMS experiments confirm that both substrates are not only adenylated but also loaded onto the T-domain.

2.2.6. Role of Putative Thioesterase Motif in PhsA

Instead of a typical Te domain at the *C*-terminus of NRPSs in the termination module (i.e. either PhsB or PhsC in PTT biosynthesis), a single thioesterase (Te) GXSXG motif is found in the *N*-terminus of PhsA (Figure 2.3*C*). In addition, two genes *theA* and *theB* were identified within the PTT biosynthetic gene cluster, and their gene products showed high similarity to type II thioesterases (TeII).[27,35] Within polyketide and nonribosomal polypeptide biosyntheses, it has been shown that TeII enzymes may have editing roles.[52,53] In order to analyze the thioesterase function of these genes, Schinko and coworkers have generated several mutants in the producing strain and showed that the thioesterase motif in the *N*-terminus of PhsA is required for PTT production. On the other hand, TheA and TheB might have editing functions as their knockout mutants only led to reduced PTT production.[35]

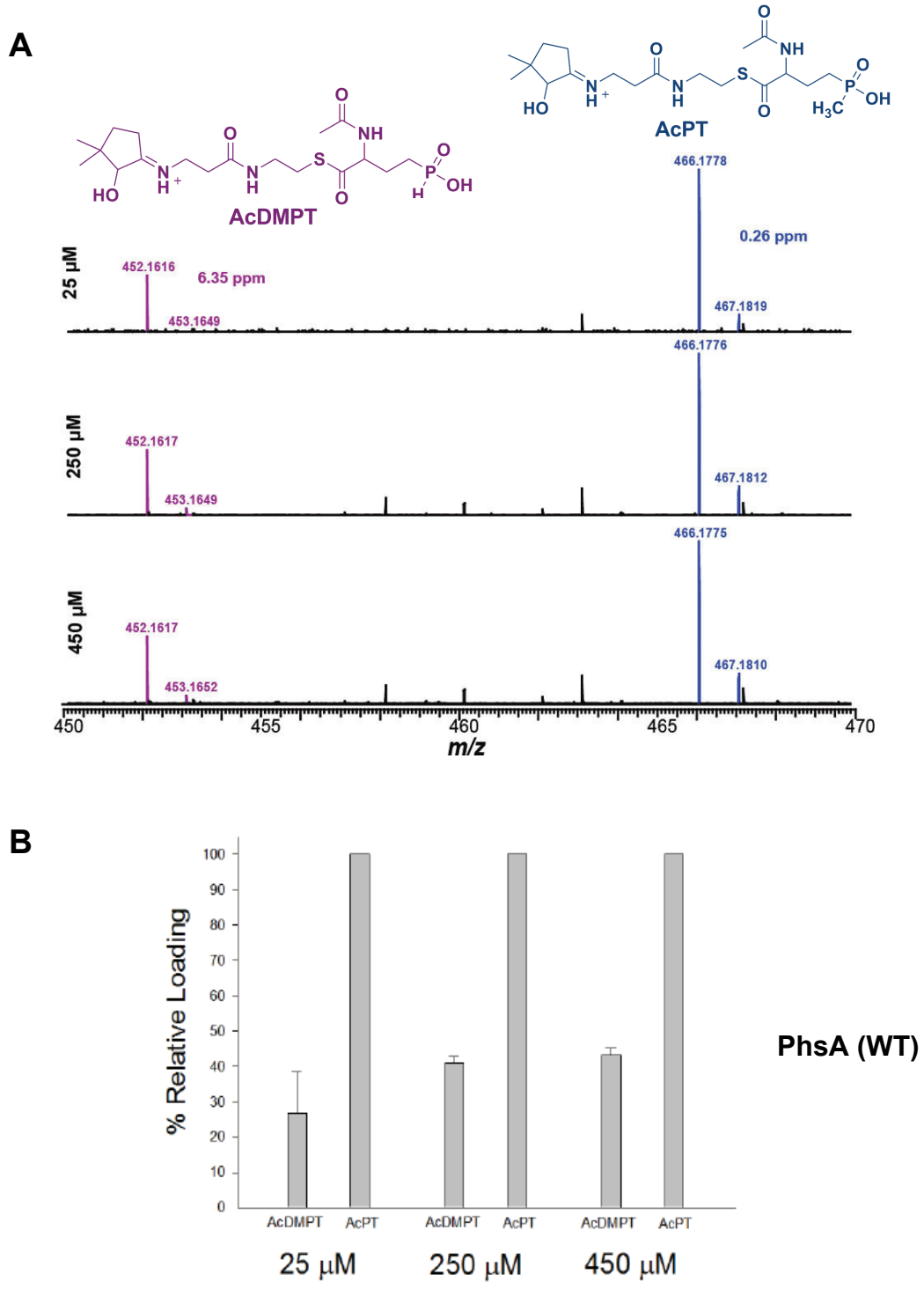


Figure 2.10 Competition loading results determined by the LC-FTMS PPant ejection assay at the indicated substrate concentrations. (A) Proposed structures of PPant ejection ions colored to match corresponding peaks. (B) Peak intensities of PPant ejection ions resulting from loaded AcPT and AcDMPT. These figures were prepared by Bradley Evans, and used with his permission.

If the Te motif in PhsA were to selectively remove AcDMPT (but not AcPT) after loading onto the PPant arm of the PCP domain, the higher ATP-PP_i exchange activity of AcDMPT at high substrate concentration would not conflict with the lower amounts of loaded AcDMPT seen in Figure 2.10*B*. Two experiments were conducted to investigate the possible involvement of this Te motif in editing. First, the putative catalytic $Ser16^{35}$ in the Te motif was mutated to Ala (MBP-PhsA-His_6 S16A). The amino acid loading experiment described above was repeated with this mutant enzyme and the relative loading onto the T-domain was determined in a identical manner as was with the wild type PhsA. As shown in Figure 2.11, AcPT remained preferentially loaded at all substrate concentrations, suggesting the Te motif is not involved in any post-loading editing. In a second complementary experiment, potential time-dependent hydrolysis of thioester-bound intermediates after loading was investigated with wild type MBP-PhsA-His_6 and compared with the S16A mutant MBP-PhsA-His_6. As shown in Figure 2.12, the ratio of AcDMPT to AcPT did not change much during a 1 h time course, suggesting that post-loading editing also does not occur in S16A mutant MBP-PhsA-His_6.

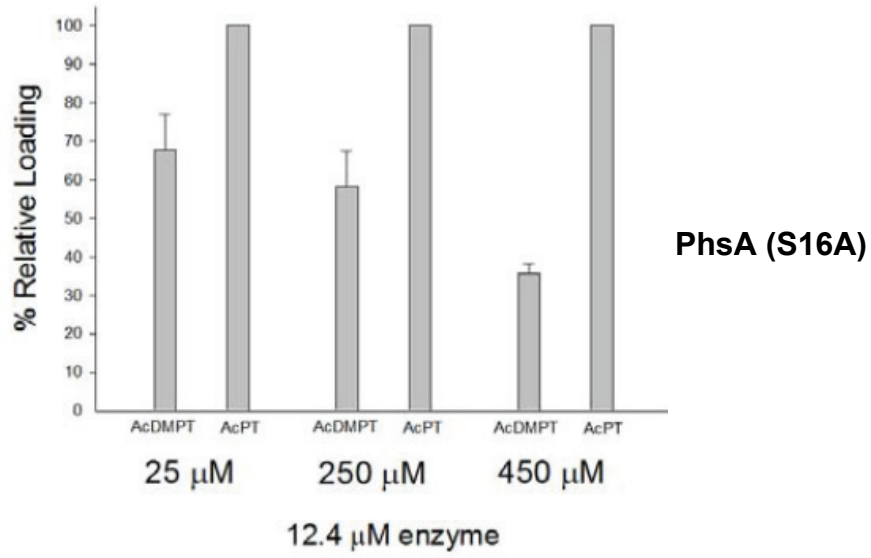


Figure 2.11 Peak intensities of PPant ejection ions resulting from loaded AcPT and AcDMPT in the assay with the MBP-PhsA(S16A)-His_6 mutant protein.

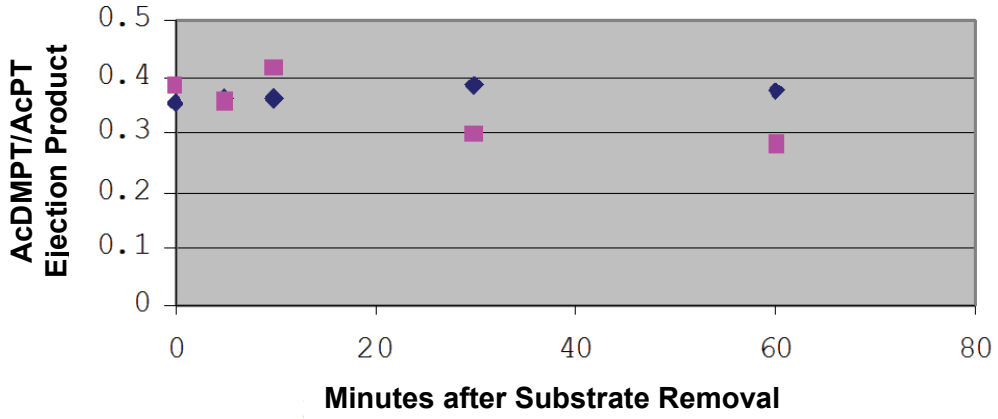


Figure 2.12 Relative peak intensities of AcDMPT to AcPT at 5, 10, 30 and 60 min after passage through a gel filtration spin column to remove small molecule substrates. WT MBP-PhsA-His_6 (◆); MBP-PhsA-His_6 S16A (■).

2.3 – Discussion and Conclusion

A high-level heterologous overexpression and purification system for PhsA was required in order to thoroughly study the first nonribosomal peptide synthetic step in PTT biosynthesis. This goal was accomplished using an MBP fusion protein, which resulted in a soluble enzyme with relative ATP-PP_i exchange activity for three substrates (AcPT, AcDMPT, and AcGlu) comparable with that reported for the native enzyme isolated from *S. viridochromogenes*. The specific activity of the recombinant enzyme for the ATP-pyrophosphate exchange activity under similar conditions is slightly (1.7-fold) higher than that of the native enzyme. The MBP tag could not be removed using Factor Xa without degradation of PhsA, suggesting that the protein is not very stable, consistent with the previous study on the native enzyme.[30]

The much lower $K_{m,app}$ for AcPT compared to that for AcDMPT may indicate that this compound is the physiological substrate for PhsA, provided that free AcDMPT and AcPT are both biosynthetic intermediates (i.e. that the P-methyltransferase PhpK acts on AcDMPT in solution, vide infra). The preference for AcPT even at high substrate concentrations where AcDMPT is a better substrate for the PP_i exchange reaction could be due to selective substrate loading with a preference for adenylated AcPT over adenylated AcDMPT, a fidelity filter not usually observed in NRPS systems.[43,44] Alternatively, the putative Te motif at the *N*-terminus of PhsA could have been involved in proofreading and catalyze selective hydrolysis of misloaded amino acids in a posttransfer editing step. However, in the test with the S16A mutant, the same selectivity was observed as for the wild-type enzyme, ruling out an editing role for the Te motif. Another possibility is that the MBP and His_6 tags in the fusion protein influence the

activity of the protein. However, these tags do not appear to affect the amino acid dependent ATP-pyrophosphate exchange reaction since the observed activities are very similar to those of the native enzyme, and it seems unlikely (although possible) that loading onto the T-domain is selectively affected by the tags.

Collectively, the observations in this study do not clarify the identity of the physiological substrate for PhsA and thus the timing of methylation on phosphorus. On one hand, a *phsA* disruption mutant accumulated AcDMPT, suggesting that this compound is the natural substrate.[36,37] On the other hand, the results reported here suggest it may be AcPT. In support of the latter model, Seto and co-workers showed that a mutant producer strain accumulated AcPT and AcPT-Ala-Ala, suggesting free AcDMPT can be methylated by PhpK.[34] Similarly, the cell-free extract of a *Streptomyces lividans* strain carrying a plasmid with a 9.5 kb fragment of the PTT biosynthetic gene cluster from *S. hygroscopicus* was able to carry out the methylation of both AcDMPT and AcDMPT-Ala-Ala.[33] Of note, this DNA fragment carries the *phpK* and *phsA* genes but not the genes for PhsB and PhsC. Since AcDMPT-Ala-Ala is unlikely to be loaded onto PhsA, these results suggest the P-methyltransferase can act on free AcDMPT-Ala-Ala (free meaning not bound to an NRPS) and presumably also on free AcDMPT. This observation in turn would be consistent with AcPT being the substrate for PhsA. In this scenario, the observed accumulation of AcDMPT upon disruption of the *phsA* gene could be the result of disruption of protein-protein interaction. Such accumulation of a biosynthetic intermediate several steps upstream of a disrupted gene is not unprecedented in secondary metabolism. In summary, this study shows successful establishment of an overexpression system for PhsA in *E. coli* which opens the door for further investigations

of PhsA activity and bioengineering of PTT derivatives. On-going experiments in collaboration with the Nair group involve the structural analysis of PhsA which may provide a better understanding of substrate binding in the active site of PhsA. Finally, the results of this study lend support to the hypothesis that the biosynthesis of PTT involves P-methylation prior to peptide assembly.

2.4 – Experimental Procedures

2.4.1. Construction of pET15b-phsA, pMAL-c2x-phsA and pET28b-malE-phsA

General Materials and Methods. All molecular biology techniques, including polymerase chain reaction (PCR), transformation and plasmid preparation, were performed using standard procedures.[54] Restriction endonucleases, T4 DNA ligase, and polymerases were purchased from either New England Biolabs (Ipswich, MA) or Invitrogen (Carlsbad, CA). Failsafe PCR Premix buffers used in the GC-rich DNA PCR were purchased from Epicentre (Madison, WI). PCR amplifications were performed using an automated thermocycler (PTC 150, MJ Research). DNA fragments used in cloning experiments were isolated by gel purification using the Promega Wizard PCR cleanup kit (Madison, WI). Plasmid isolation was accomplished by Qiagen's Midiprep or Miniprep kits (Valencia, CA). Chemically competent *E. coli* DH5α cells were purchased from the University of Illinois at Urbana-Champaign (UIUC) Cell Media Facility. Sequencing reactions were performed at the W.M. Keck Center for Biotechnology at the University of Illinois at Urbana-Champaign utilizing standard T7 forward and reverse primers (T7-for/T7-rev). The PCR products were digested with the appropriate restriction endonucleases and then ligated into vector DNA that was digested with the same

restriction enzymes. Chemically competent *E. coli* DH5α cells were transformed with each ligation mixture and plated on Luria-Bertani (LB)-agar containing the appropriate antibiotic to select for positive clones. For pET15b and pMAL-c2x-based constructs, ampicillin (Sigma-Aldrich) was used at 100 μg/mL, and for pET28b-based constructs – kanamycin (Fisher) was used at 50 μg/mL concentration.

pET15b-phsA [JHL-I-48-53, 59-62, 64, 77-80, 82-89]. The genomic DNA of *S. viridochromogenes* DSM 40736 was prepared in the laboratory of Professor William W. Metcalf (Department of Microbiology, UIUC). From this genomic DNA the *phsA* gene was amplified by PCR using a forward primer 5'-GGGAATTCCATATGACCGCAGCG ACACCGGACA–3'containing an *Nde*I restriction site (underlined) and a reverse primer 5'-TATAAACGCGGATCCCTACGTCCCCTTCAGTTC–3' containing a *Bam*HI restriction site. Successful PCR products were digested with *Nde*I and *BamH*I and subsequently ligated into the double-digested pET15b vector (Novagen) with T4 DNA ligase.

pMAL-c2x-phsA [JHL-II-45-79]. The *phsA* gene was amplified by PCR as described above using a forward primer 5'-GGAAGGATTTCAGAATTCATGACCGCAGCG-3' containing an *Eco*RI site and a reverse primer 5'-TATAAACGCGGATCCCTACGTCCC CTTCAGTTC-3' containing a *Bam*HI site. The PCR product and pMAL-c2x vector were treated with the corresponding restriction enzymes and subjected to double ligation to produce a vector encoding PhsA fused at its *N*-terminus to maltose binding protein.

pET28b-malE-phsA [JHL-III-45-56, 58-64, 70-76]. The *malE-phsA* gene was amplified by PCR from the pMAL-c2x-*phsA* construct using a forward primer 5'-GCCGCTCGCCCATGGCTATGAA AATCGAAGAAGGTAAACTGG-3' and a reverse primer 5'-AGAGCGAAGAGCCTCGAGC GTCCCCTTCAGTTCGCG-3' to introduce *Nco*I and *Xho*I restriction sites, respectively. The PCR product and pET28b vector were treated with these enzymes and then ligated to produce a vector encoding MBP-PhsA with a hexahistidine tag at its *C*-terminus.

pET28b-malE-phsA (S16A) [JHL-VII-95-99, JHL-IX-11-20, 27]. The single point mutation encoding construct, pET28b-malE-phsA (S16A) mutant could not be prepared by a simple Quick-Change method, instead the "MegaWhop" method[55] was used in which mega primers of more than 1.5 kb in length were used for whole plasmid PCR reaction. The mega primer was prepared by PCR from pET28b-malE-*phsA* wild type template using a forward primer 5'-CCGGACCGCACCGGCCCGGCCCCGGGGGCCT GCCCCGTC-3' which introduces S16A mutant coding gene and a reverse primer 5'-TGAGCGGCCTGGGCC-3'. DpnI was used to digest the methylated wild-type template.

2.4.2. Expression and Purification of PhsA with Various Tags

Expression of His_6-PhsA [JHL-I-90-98, JHL-II-7-23, 32-38]. Electrocompetent *E. coli* BL21 (DE3) cells were prepared using standard methods.[54] The cells were freshly transformed with pET15b-*phsA*. Single colony transformants were grown aerobically for 12-15 h at 37 °C in 5 mL of LB medium supplemented with 100 μg/mL ampicillin. The fully grown cell culture was centrifuged at 5000×*g* for 15 min and the spent medium was

discarded. The pellet was carefully resuspended in fresh LB medium and the resuspended cells were used to inoculate a large-scale culture of 2-3 L of LB-ampicillin (100 μg/mL). The culture was grown with aeration at 37 °C until the optical density at 600 nm (OD_{600}) reached 0.5-0.7. The expression of *N*-terminally His_6-tagged PhsA was induced by addition of isopropyl-1-thio-*β*-D-galactoside (IPTG) from CalBiochem to a final concentration of 0.1 mM and the culture was continually shaken at 18 °C for an additional 12-15 h. After harvesting the cells at 5000×*g* for 15 min, the spent medium was discarded and the cell pellet was resuspended in buffer A (50 mM Tris, 300 mM NaCl, 15% glycerol (v/v), 10 mM β-mercaptoethanol, pH 7.8 at 4 °C) and lysed by incubation with lysozyme at a final concentration of 0.2 mg/mL for 30 min at RT before being analyzed for solubility of the expressed protein.

Expression and Purification of MBP-PhsA [JHL-II-83-88, 91-92, JHL-III-6, 17, 23-24, 26-41]. Electrocompetent *E. coli* BL21 (DE3) and Rosetta 2 (DE3) cells were prepared using standard methods,[54] and used as expression hosts. Transformations of these strains with pMAL-c2x-*phsA* were carried out in a similar manner as described for the expression of His_6-PhsA, except in the case of Rosetta 2 (DE3), a chloramphenicol (20 μg/mL) marker was used in addition to ampicillin (100 μg/mL). The overnight starter culture was used to inoculate the 2 - 3 L of LB supplemented with the appropriate antibiotics as well as glucose (0.2% w/v final concentration) in order to repress expression of the maltose genes of the *E. coli*. The culture was grown at 37 °C until the optical density of A_{600} reached 0.5-0.7. The expression of PhsA as *N*-terminal MBP-fusion was induced by addition of IPTG to a final concentration of 0.5 mM and the

culture was continued to grow at 18 °C for an additional 12-15 h. The cells were harvested at 5000×g for 15 min, the spent medium discarded, and the cell pellet resuspended in buffer A, supplemented with protease inhibitor cocktail tablets (Complete Mini, EDTA-free, Roche).

Cell lysis was achieved by incubation of the cells in suspension buffer with lysozyme at a final concentration of 0.2 mg/mL for 30 min at RT followed by sonication (Sonics & Materials Inc. Vibra Cell^{TM}, 35% Amp, pulse sequence of 5.0 s on and 9.9 s off, 15 min) in an ice-water bath. The lysate was cleared by centrifugation (30 min, 25,000×*g*), and passed through a 0.45 μm syringe filter (Corning) onto the amylose resin, which had been equilibrated with buffer B (20 mM Tris, 200 mM NaCl, pH 7.4 at 4 °C). After washing the protein-bound amylose resin with 8-10 column volumes of buffer B, the MBP-fusion protein was eluted from the column by elution with buffer C (20 mM Tris, 200 mM NaCl, 10 mM maltose, pH 7.4 at 4 °C). The elution fractions that contain the desired MBP-fusion protein as analyzed by SDS-PAGE (10% acrylamide gel in Tris-glycine buffer) were combined and concentrated using a centrifugal filtering device (Ultra Amicon, Millipore, 30 kDa molecular weight cutoff). After excess NaCl and maltose were removed using a PD-10 desalting column (GE Healthcare) pre-packed with Sephadex G-25 medium, the protein was eluted with buffer D (20 mM Tris, 25 mM NaCl, pH 7.5). The purified MBP-PhsA was treated with Factor Xa protease to cleave MBP from the fusion protein (described below), and PhsA was separated from MPB by DEAE-sepharose anion exchange chromatography with increasing NaCl concentration (from 25 mM up to 500 mM) in the 20 mM Tris buffer. All purification steps were performed at 4 °C. The protein concentration was determined using A_{280} measurement as well as

Bradford assay with Bovine Serum Albumin (BSA) as protein standard.

To improve protein folding and thereby reduce possible proteolytic degradation of MBP-PhsA during purification as well as during the Factor Xa cleavage reaction, chaperone proteins were co-expressed with MBP-PhsA in a commercially available host, BL21-pGKJE8 (Takara). After transformation of the strain by the method described above, a single colony was used to generate an overnight culture, which was then used to inoculate 3 L of LB-glucose medium supplemented with chloramphenicol (20 μg/mL) and ampicillin (50 μg/mL). The chaperone proteins, Dnak, DnaJ and GrpE were induced by L-arabinose (0.5 mg/mL, Aldrich), and GroEL and GroES were induced by tetracycline (5 ng/mL, Aldrich) both at the onset of the cell growth. When OD_{600} reached between 0.5 – 0.7, the cells were induced by addition of IPTG at a final concentration of 0.5 mM. The culture was continued to grow at 37 °C for another 4 h or at 25 °C for 6 h, and harvested at 5000×g for 15 min.

Expression and Purification of MBP-PhsA-His_6 [JHL-III-68-69, 78-80, 84-85, JHL-IV-6, 37, 53, JHL-V-10-11, JHL-VI-25, 87]. BL21-pGKJE8 strain was transformed with pET28b-*malE-phsA*, and following the procedure described above, the culture was grown at 37 °C in 3 L of LB medium supplemented with chloramphenicol (20 μg/mL) and kanamycin (50 μg/mL) as selection markers. Both L-arabinose (0.5 mg/mL) and tetracycline (5 ng/mL) were supplemented to induce the chaperone proteins. When the OD_{600} reached between 0.5 - 0.7, the cell culture was induced with (IPTG) to a final concentration of 0.1 mM, and continued to grow at 18 °C for an additional 12-15 h. The cells were harvested by centrifugation (15 min, 5000g), resuspended in buffer A with

protease inhibitor cocktail tablets (Complete Mini, EDTA-free, Roche), and disrupted by incubation with 0.2 mg/mL of lysozyme followed by sonication (pulse sequence of 5 s on and 9.9 s off, 15 min). The lysate was cleared by centrifugation (30 min, 30,000×*g*), and purified using Ni^{2+} affinity resin (Qiagen) equilibrated with buffer A containing 10 mM imidazole. The cell lysate was passed through a 45 μm filter before application to the resin. The column (1 x 10 cm) was washed with 20 column volumes of buffer A containing 25 mM imidazole. Elution of the desired His_6-tagged protein was achieved using buffer A with increasing imidazole concentrations from 25 mM to 500 mM in 25 mM increments (total volume = 200 mL), and the fractions containing the desired protein as determined by SDS-PAGE and Coomassie Blue staining were concentrated using Amicon Ultra with a 30 kDa membrane. Imidazole and NaCl were removed using a PD-10 desalting column. The protein was eluted with buffer D, and was further purified on a HiTrap Q HP anion exchange column (GE Healthcare) pre-equilibrated with buffer B. Elution of the enzyme was carried out with buffer B in a step gradient of NaCl, from 25 mM to 500 mM with 25 mM increments (total volume = 220 mL). The purified protein MBP-PhsA-His_6 was concentrated using Amicon Ultra (30 kDa), mixed with glycerol to a final concentration of 20% (w/v), flash frozen in liquid nitrogen, and stored at -80 °C. Protein concentration was determined by Bradford assay using BSA as a protein standard.

Factor Xa Cleavage of the MBP-fusion protein [JHL-II-94-95, JHL-III-6, 17, 34, JHL-IV-12, 16]. Either MBP-PhsA or MBP-PhsA-His_6 was incubated with Factor Xa (New England Biolabs) at 1% w/w (Factor Xa protease per total protein) in the elution buffer from the purification, typically for 3 h at either 4 °C or 25 °C.

2.4.3. In Vitro Phosphopantetheinylation of MBP-PhsA-His$_6$ by Sfp

General Materials. ^{3}H-Acetyl coenzyme A ([^{3}H]CoA) was purchased from MP Biomedicals, and radioactive tetrasodium ^{32}P-pyrophosphate was purchased from Perkin Elmer. Both Sfp and BODIPY-CoA were prepared by the Kelleher group (Department of Chemistry, UIUC).

Conversion of Apo-MBP-PhsA-His$_6$ to Holo-MBP-PhsA-His$_6$ [JHL-IV-10, JHL-V-61, JHL-VI-46-51]. For *in vitro* priming of MBP-PhsA-His$_6$ to its holo form, 9.6 μM MBP-PhsA-His$_6$ was incubated with 25 μM CoA and 1 nmol Sfp for 1 h at 25 ^{o}C in assay buffer (50 mM HEPES pH 7.5, 1 mM tris(2-carboxyethyl)phosphine (TCEP), 10 mM $MgCl_2$) in a final volume of 150 μL.

In-Gel Fluorescence Labeling Using BODIPY-CoA. MBP-PhsA-His$_6$ was incubated with Sfp (4 μM) in a buffer containing 10 mM $MgCl_2$, 2.4 mM DTT and 0.16 mM BODIPY-CoA for 1-2 h at 30 °C. The reaction was quenched by adding an equal volume of SDS-PAGE loading buffer. The reaction mixture was analyzed by running a 10% SDS-PAGE gel, exciting the fluorophore at 460 nm, and detecting at 520 nm. Subsequently, the gel was stained with Coomassie blue.

Radioassay of the Phosphopantetheinylation Process [JHL-VI-55-59]. The activity of Sfp on MBP-PhsA-His$_6$ was also monitored by measuring the incorporation of tritium-labeled acetyl CoA into apo-MBP-PhsA-His$_6$ (13 μM) under the phosphopantetheinylation

conditions described above except for the use of [^{3}H]CoA (20 mCi/mmol, 10 μCi/mL, 52 μM) in place of CoA and the use of a lower Sfp concentration (0.1 μM). Reactions were initiated by addition of Sfp, and at each time-point, a 30 μL aliquot was taken from the reaction mixture, and quenched with 240 μL of 10% trichloroacetic acid (TCA). BSA (100 μg) was added as a carrier protein, and the precipitated protein was pelleted by centrifugation. After the supernatant was discarded, the pellet was washed three times with 240 μL of 10% TCA before resuspension in 100 μL formic acid. Liquid scintillation cocktail (ScientiSafe Econo 1, Fisher Scientific) was added to the protein suspension, and the radioactivity of tritium bound to the protein was quantified by liquid scintillation counting.

2.4.4. Determination of Substrate Specificities

Kinetic Analysis of the Pyrophosphate Exchange Reaction [JHL-IV-19-24, 42, 56-60, 85-98, JHL-V-15-17, 27-32, 40, 44-52, 64, JHL-VI-21-22, 46-51]. The reactions were carried out in 90 μL total volume containing 50 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 2 mM Na_4PP_i, 1 μCi [^{32}P]PP_i, 0.005 - 4 mM of substrate and 0.52 μM of holo MBP-PhsA-His_6. The reaction mixture was equilibrated at 30 °C before initiation of the reaction with addition of 2 mM ATP. Aliquots were taken at each time-point, and analyzed by thin layer chromatography (TLC) using polyethyleneimine plates (Scientific Adsorbents Inc., Atlanta, GA) pre-run in water. ATP and pyrophosphate were resolved using a developing solvent of 750 mM KH_2PO_4, pH 3.5, and 4 M urea. Separated radiolabeled products were visualized using a phosphorimaging plate (Fujifilm, BAS-MS2040), and the scanned image (Storm^{TM}) was analyzed using ImageQuant software. In

addition to the expected exchange of ^{32}P from pyrophosphate into ATP, small amounts of radiolabeled ADP and trace amounts of P_i were also observed, as has been seen in other studies of adenylation domains.[56] ADP is typically not identified using the charcoal adsorption method as both ATP and ADP are adsorbed to the charcoal. The formation of these additional products was dependent on both PhsA and substrate. Inclusion of these radiolabeled products resulted in the same trends (larger apparent k_{cat} for AcDMPT, smaller apparent K_m for AcPT) as did omitting them from the analysis. For apo-MBP-PhsA-His_6, the protein was treated under the same conditions as for holo protein.

FTMS Analysis of MBP-PhsA-His_6 after Loading with AcPT and AcDMPT [JHL-VI-75, JHL-VII-10]. Samples for FTMS were prepared in the same manner as in the pyrophosphate exchange assay except that higher concentrations of protein were utilized. A solution of 46.5 μM MBP-PhsA-His_6 in 70 mM Tris-HCl, 10 mM $MgCl_2$, 2 mM DTT, 4 μM Sfp, and 200 μM CoA was incubated at 30 ^oC for 30 min for phosphopantetheinylation. To the solution was then added ATP (10 mM final concentration) and 25, 250 or 450 μM AcDMPT or AcPT (final concentrations) and the assays were incubated for 1 h at 30 ^oC followed by quenching with 5% formic acid, centrifugation to remove any particulates, and analysis by LC-MS. All mass spectral data were collected by Bradley S. Evans (Kelleher group) on a 7 Tesla LTQ-FT (Thermo Fisher Scientific) equipped with a Surveyor MS pump and autosampler. Reverse phase liquid chromatography (RPLC) was carried out on a Jupiter 2.0 X 50 mm C4 column (Phenomenex) using water (solvent A) and acetonitrile (solvent B) with 0.1% formic acid. LC-MS assays used a gradient program starting at 30% B increasing to 65% B over 15

min followed by a 20 min re-equilibration to initial conditions. The mass spectrometric method consisted of a low resolution ion trap full MS scan followed by a pseudo-MS^2 scan (NS at 75 V) for the phosphopantetheine (PPant) ejection assay. Note that all species entering the mass spectrometer are fragmented simultaneously giving rise to highly complicated MS/MS data; however, the <5 part-per-million mass accuracy is the key to monitoring the specific m/z channels corresponding to the ions from PPant ejection. LC-MS data was analyzed using QualBrowser software (Thermo Fisher). For quantitation of PPant ejection products, Nozzle-Skimmer dissociation (NS) scans were summed over the period of MBP-PhsA-His_6 elution for each LC-MS run to obtain the total normalized levels of the PPant ejection ions. The predominant product was set to 100% and the substrate with lower activity was normalized to that value. Results from five replicate assays were used for comparison, with an observed precision of 9.7, 7.3 and 4.1% at 1 sigma for the 25, 250 and 450 μM assays, respectively.

2.4.5. Synthesis of Substrate Analogues

General Materials and Methods. All NMR spectra were recorded on Varian U400 or Varian U500 spectrometers. ^1H NMR spectra were referenced to TMS at 0 ppm or $CHCl_3$ at 7.26 ppm, and ^{13}C NMR spectra were referenced to $CDCl_3$ at 77.7 ppm. Mass spectrometry (MS) experiments were carried out by the Mass Spectrometry Laboratory at the University of Illinois at Urbana-Champaign. Fractions collected during silica gel column chromatography were analyzed by TLC. Unless otherwise specified, all compounds and solvents were obtained from Fisher or Aldrich. THF was distilled over sodium/benzophenone, and CH_2Cl_2 was distilled over CaH_2 prior to use. AcDMPT was

synthesized and characterized as described in the Supporting Information of Reference (31). L-(+)-2-Amino-4-phosphonobutyric acid (L-AP4) was purchased from Tocris Bioscience. Both L-Glu and L-Asp were purchased from Sigma-Aldrich, and PT was obtained from Matrix Scientific. These acids were acetylated with acetic anhydride in the presence of potassium carbonate in water *[JHL-III-91-93, JHL-IV-13-15]*. Substrate concentrations in stock solutions were determined by integration of the corresponding phosphorus peak area in the ^{31}P NMR spectra.

2-Aminobutyrolactone hydrobromide (**1**) *[JHL-I-4-6, 8, 27, 37]*.[57] To a 100 mL round-bottomed flask containing L-methionine (2.14 g, 14.3 mmol) was added 10 mL of H_2O, 10 mL of isopropyl alcohol, and 4 mL of glacial acetic acid. To this suspension was added bromoacetic acid (18.64 mmol) and the reaction was stirred under N_2 at 50 °C until all solids had dissolved. The solution was then refluxed at 100-110 °C for 6 h. The solvent was removed under reduced pressure using a rotary evaporator and the resulting oil was redissolved in 1.5 mL of 1:1 of isopropyl alcohol/toluene. The suspension was concentrated under reduced pressure to remove remaining H_2O giving a yellow emulsion. To this, 9 mL of 4 M HCl in dioxane was added, stirred at room temperature for 20 min, and the reaction was placed at 4 °C overnight. The solid was filtered and washed with cold dioxane to produce 1.9 g of pinkish solid product (73-81%), m.p. 221–224 °C (lit. 234–236 °C).[21] ^{1}H NMR (500 MHz, D_2O) δ 4.45 (t, 1H); 4.32-4.28 (m, 2H); 2.67-2.61 (m, 1H); 2.28 (q, 1H), ^{13}C NMR (500 MHz, D_2O) δ 174.67, 67.44, 48.61, 26.85.

Aminobromobutyrate hydrobromide (**2**) *[JHL-I-7, 17, 28, 65]*. A high pressure glass tube was charged with 2 g of compound **1** and 11 mL of 33% HBr/HOAc. The suspension was stirred at 35 ˚C for 30 min, and then the temperature was raised to 75 ˚C. After 20 h, a white solid of bromide was obtained, which was filtered and washed with cold ether to yield white product as a solid (94%), m.p. 181-183 °C (lit 187-188 °C).[37] ^{1}H NMR (500 MHz, D_2O) δ 4.09 (t, 1H); 3.45 (m, 2H); 2.45-2.22 (m, 2H).

Aminobromomethyl butanoate (**3**) *[JHL-I-9, 18, 29, 66]*. Acetyl chloride (14 mL, 10 equiv.) was added dropwise to anhydrous MeOH (100 mL) at 0 ˚C. Upon completion of addition, the mixture was allowed to warm to room temperature and was stirred for 30 min before 5.4 g of compound **2** (20.5 mmol) was added. The mixture was then stirred for another 20 h. The solvent was removed under reduced pressure to afford methyl ester as yellow oil in a quantitative yield. ^{1}H NMR (500 MHz, D_2O) δ 4.18 (t, 1H); 3.68 (s, 3H); 3.45 (m, 2H); 2.45-2.22 (m, 2H). ^{31}C NMR is as reported in the literature.[37]

2.5 – References

1. Seto, H.; Kuzuyama, T.: "Bioactive Natural Products with Carbon-Phosphorus Bonds and Their Biosynthesis." *Nat. Prod. Rep.* **1999**, *16*, 589-596.

2. Diddens, H.; Zahner, H.; Kraas, E.; Gohring, W.; Jung, G.: "Transport of tripeptide antibiotics in bacteria." *Eur. J. Biochem.* **1976**, *66*, 11-23.

3. Bayer, E.; Zahner, H.; Konig, W. A.; Jessipow, S.; Gugel, K. H.; Hagele, K.; Hagenmai.H: "Metabolites of microorganisms 98. Phosphinothricin and phosphinothricyl alanylalanine." *Helv. Chim. Acta* **1972**, *55*, 224.

4. Strauch, E.; Wohlleben, W.; Puhler, A.: "Cloning of a phosphinothricin *N*-acetyltransferase gene from *Streptomyces viridochromogenes* Tü94 and its expression in *Streptomyces lividans* and *Escherichia coli.*" *Gene* **1988**, *63*, 65-74.

5. Kumada, Y.; Anzai, H.; Takano, E.; Murakami, T.; Hara, O.; Itoh, R.; Imai, S.; Satoh, A.; Nagaoka, K.: "The bialaphos resistance gene (*bar*) plays a role in both self-defense and bialaphos biosynthesis in *Streptomyces hygroscopicus.*" *J. Antibiot.* **1988**, *41*, 1838-1845.

6. Schwartz, D.; Berger, S.; Heinzelmann, E.; Muschko, K.; Welzel, K.; Wohlleben, W.: "Biosynthetic gene cluster of the herbicide phosphinothricin tripeptide from *Streptomyces viridochromogenes* Tü94." *Appl. Environ. Microbiol.* **2004**, *70*, 7093-7102.

7. Blodgett, J. A. V.; Zhang, J. K.; Metcalf, W. W.: "Molecular cloning, sequence analysis, and heterologous expression of the phosphinothricin tripeptide biosynthetic gene cluster from *Streptomyces viridochromogenes* DSM 40736." *Antimicrob. Agents Chemother.* **2005**, *49*, 230-240.

8. Hara, O.; Murakami, T.; Imai, S.; Anzai, H.; Itoh, R.; Kumada, Y.; Takano, E.; Satoh, E.; Satoh, A.; Nagaoka, K. "The bialaphos biosynthetic genes of *Streptomyces viridochromogenes:* Cloning, heterospecific expression, and comparison with the genes of *Streptomyces hygroscopicus." J. Gen. Microbiol.* **1991**, *137*, 351-359.

9. Murakami, T.; Anzai, H.; Imai, S.; Satoh, A.; Nagaoka, K.; Thompson, C. J.: "The bialaphos biosynthetic genes of *Streptomyces hygroscopicus:* Molecular cloning and characterization of the gene cluster." *Mol. Gen. Genet.* **1986**, *205,* 42-50.

10. Grammel, N.; Schwartz, D.; Wohlleben, W.; Keller, U.: "Phosphinothricin Tripeptide synthetases from *Streptomyces viridochromogenes." Biochemistry* **1998**, *37,* 1596-1603.

11. Schwartz, D.; Alijah, R.; Nussbaumer, B.; Pelzer, S.; Wohlleben, W.: "The Peptide synthetase gene *phsA* from *Streptomyces viridochromogenes* is not juxtaposed with other genes involved in nonribosomal biosynthesis of peptides." *Appl. Environ. Microbiol.* **1996**, *62*, 570-577.

12. Schwartz, D.; Grammel, N.; Heinzelmann, E.; Keller, U.; Wohlleben, W.: "Phosphinothricin tripeptide synthetases in *Streptomyces viridochromogenes* Tü94." *Antimicrob. Agents Chemother.* **2005**, ***49***, 4598-4607.

13. Kamigiri, K.; Hidaka, T.; Imai, S.; Murakami, T.; Seto, H.: "Studies on the biosynthesis of bialaphos (SF-1293) .12. C-P bond formation mechanism of bialaphos: discovery of a P-methylation enzyme." *J. Antibiot.* **1992**, *45*, 781-787.

14. Imai, S.; Seto, H.; Sasaki, T.; Tsuruoka, T.; Ogawa, H.; Satoh, A.; Inouye, S.; Niida, T.; Otake, N.: "Studies on the biosynthesis of bialaphos (SF-1293) .6. Production of *N*-acetyldemethylphosphinothricin and *N*-acetylbialaphos by blocked mutants of *Streptomyces hygroscopicus* SF-1293 and their roles in the biosynthesis of bialaphos." *J. Antibiot.* **1985**, *38*, 687-690.

15. Eys, S.; Schwartz, D.; Wohlleben, W.; Schinko, E.: "Three thioesterases are involved in the biosynthesis of phosphinothricin tripeptide in *Streptomyces viridochromogenes* Tü94." *Antimicrob. Agents Chemother.* **2008**, *52*, 1686-1696.

16. Alijah, R.; Dorendorf, J.; Talay, S.; Puhler, A.; Wohlleben, W.: "Genetic analysis of the phosphinothricin tripeptide biosynthetic pathway of *Streptomyces viridochromogenes* Tü94." *Appl. Microbiol. Biotechnol*. **1991**, *34*, 749-755.

17. Wohlleben, W.; Alijah, R.; Dorendorf, J.; Hillemann, D.; Nussbaumer, B.; Pelzer, S.: "Identification and characterization of phosphinothricin tripeptide biosynthetic genes in *Streptomyces viridochromogenes."* Gene **1992**, *115*, 127-132.

18. Cole, P. A.: "Chaperone-assisted protein expression." *Structure* **1996**, *4*, 239-242.

19. Issleib, K.; Mogelin, W.; Balszuweit, A.: "Bis(trimethylsilyl) hypophosphite and alkoxycarbonylphosphonous acid bis(trimethylsilyl) esters as building-blocks in organo-phosphorus chemistry." *Z. Anorg. Allg. Chem*. **1985**, *530*, 16-28.

20. Deprele, S.; Montchamp, J. L.: "Triethylborane-initiated room temperature radical addition of hypophosphites to olefins: synthesis of monosubstituted phosphinic acids and esters." *J. Org. Chem.* **2001**, *66*, 6745-6755.

21. Vu, M. T.; Martinis, S. A.: "A unique insert of leucyl-tRNA synthetase is required for aminoacylation and not amino acid editing." *Biochemistry* **2007**, *46*, 5170-5176.

22. Cole, F. X.; Schimmel, P. R.: "On rate law and mechanism of adenosine triphosphate-pyrosphosphate isotope exchange reaction of amino acyl transfer ribonucleic acid synthetases." *Biochemistry* **1970***, 9*, 480.

23. Luo, L. S.; Burkart, M. D.; Stachelhaus, T.; Walsh, C. T.: "Substrate recognition and selection by the initiation module PheATE of Gramicidin S synthetase." *J. Am. Chem. Soc. 2001,* **123**, 11208-11218.

24. Keating, T. A.; Suo, Z. C.; Ehmann, D. E.; Walsh, C. T.: "Selectivity of the Yersiniabactin synthetase adenylation domain in the two-step process of amino acid activation and transfer to a hole-carrier protein domain." *Biochemistry* **2000**, *39,* 2297-2306.

25. Gatto, G. J.; McLoughlin, S. M.; Kelleher, N. L.; Walsh, C. T.: "Elucidating the substrate specificity and condensation domain activity of FkbP, the FK520 Pipecolate-incorporating enzyme." *Biochemistry* **2005**, *44*, 5993-6002.

26. Eppelmann, K.; Stachelhaus, T.; Marahiel, M. A.: "Exploitation of the selectivity-conferring code of nonribosomal peptide synthetases for the rational design of novel peptide antibiotics." *Biochemistry* **2002**, *41*, 9718-9726.

27. Baldwin, J. E.; Shiau, C. Y.; Byford, M. F.; Schofield, C. J.: "Substrate-specificity of L-delta-(alpha-aminoadipoyl)-L-cysteinyl-D-valine synthetase from Cephalosporium-Acremonium - demonstration of the structure of several unnatural tripeptide products." *Biochem. J.* **1994**, *301*, 367-372.

28. Stachelhaus, T.; Marahiel, M. A.: "Modular structure of peptide synthetases revealed by dissection of the multifunctional enzyme GrsA." *J. Biol. Chem.* **1995**, *270*, 6163-6169.

29. Dorrestein, P. C.; Bumpus, S. B.; Calderone, C. T.; Garneau-Tsodikova, S.; Aron, Z. D.; Straight, P. D.; Kolter, R.; Walsh, C. T.; Kelleher, N. L.: "Facile detection of acyl and peptidyl intermediates on thiotemplate carrier domains via phosphopantetheinyl elimination reactions during tandem mass spectrometry." *Biochemistry* **2006**, *45*, 12756-12766.

30. Hansen, D. B.; Bumpus, S. B.; Aron, Z. D.; Kelleher, N. L.; Walsh, C. T.: "The loading module of mycosubtilin: An adenylation domain with fatty acid selectivity." *J. Am. Chem. Soc.* **2007**, *129*, 6366.

31. Lee, J. H.; Evans, B. S.; Li, G. Y.; Kelleher, N. L.; van der Donk, W. A.: "In vitro characterization of a heterologously expressed nonribosomal peptide synthetase involved in phosphinothricin tripeptide biosynthesis." *Biochemistry* **2009**, *48*, 5054-5056.

32. Heathcole, M. L.; Staunton, J.; Leadlay, P. F.: "Role of type II thioesterases: evidence for removal of short acyl chains produced by aberrant decarboxylation of chain extender units." *Chem*. *Biol.* **2001***, 8*, 207-220.

33. Schwarzer, D.; Mootz, H. D.; Linne, U.; Marahiel, M. A.: "Regeneration of misprimed nonribosomal peptide synthetases by type II thioesterases." *Proc. Nat. Acad. Sci. U.S.A.* **2002***, 99*, 14083-14088.

34. Sambrook, J. F., E. F.; Maniatis, T.: *Molecular Cloning: a Laboratory Manual edn* 2nd ed. New York: Cold Spring Harbor Laboratory: Cold Spring Harbor; **1989**.

35. Miyazaki, K.; Takenouchi, M.: "Creating random mutagenesis libraries using megaprimer PCR of whole Plasmid." *Biotechniques* **2002**, *33*, 1033.

36. Rapaport, E.; Remy, P.; Kleinkauf, H.; Vater, J.; Zamecnik, P. C.: "Aminoacyl-tRNA-synthetases catalyze AMP-----ADP-----ATP exchange-reactions, indicating labile covalent enzyme-amino acid intermediates." *Proc. Nat. Acad. Sci. U.S.A.* **1987***, 84*, 7891-7895.

37. Koch, T.; Buchardt, O.: "Synthesis of L-(+)-selenomethionine." *Synthesis* **1993**, 1065-1067.

CHAPTER 3: ELUCIDATING THE BIOSYNTHETIC PATHWAY OF DEHYDROPHOS USING BIOCHEMICAL APPROACHES*

3.1 – Introduction

Dehydrophos, originally designated A53868, is a broad-spectrum antibiotic produced by *Streptomyces luridus*.[1] It was first discovered by Eli Lilly and Company in 1984, and has been shown to inhibit the growth of Gram-negative and Gram-positive bacteria and demonstrated *in vivo* activity against *Salmonella*.[1] The chemical structure of dehydrophos has been revised several times[2] until finally the correct structure was confirmed using NMR spectroscopic data and comparison to an authentic synthetic compound.[3] The revised structure of dehydrophos contains a unique *O*-methyl aminovinylphosphonate ester that is attached to the *C*-terminus of a Gly-L-Leu dipeptide (Figure 3.1).[3] Due to the structural similarity to dehydroalanine, the compound was renamed as dehydrophos. As is commonly observed in other oligopeptide phosphonate antibiotics,[4] dehydrophos is thought to be taken up by target organisms via non-specific oligopeptide permeases.[5,6] Once inside the cell, the peptide is likely hydrolyzed by peptidases to release the active phosphonate-containing part of the compound.

The important properties of C-P compounds and the common strategies found in the biosynthesis of these compounds were discussed in Chapter 1. The molecular target of dehydrophos is currently not known. Nonetheless, the free phosphonate released from dehydrophos after its hydrolysis by peptidase(s) is projected to be methyl acetylphosphonate (Figure 3.1),[7] which can potently inhibit pyruvate dehydrogenase

* Reproduced in part with permission from "Molecular Cloning and Heterologous Expression of the Dehydrophos Biosynthetic Gene Cluster" *Chem. Biol.* **2010**, *17*, 402-411. Copyright 2010 by Elsevier Ltd.

activity,[8] a crucial enzyme reaction in bacterial metabolism. Alternatively, the reactive vinyl moiety of the tripeptide could inhibit a cellular process directly, as is found in the phosphonopeptide K-26, where hydrolysis of peptide bond is not necessary for biological activity.[9] Another interesting aspect of the dehydrophos structure is the methylation of the phosphonate group. The rationale for installation of the methyl group may be the enhancement of bioavailability of dehydrophos by reducing its charge state in the ester form. Alternatively, the methyl ester may enhance interaction with the target of dehydrophos. In fact, it has been reported that methyl acetylphosphonate inhibits pyruvate dehydrogenase over 100-fold more potently than the unesterified acetylphosphonate.[8] The methylation reaction of the phosphonic acid group during dehydrophos biosynthesis will be discussed in depth in Chapter 4.

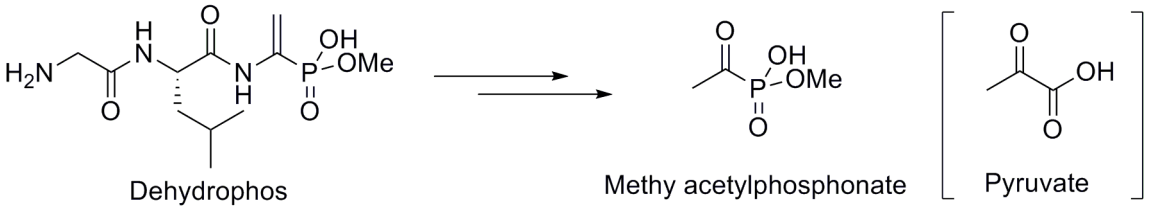


Figure 3.1 The confirmed chemical structure of dehydrophos. The compound is made of a tripeptide, Gly-L-Leu-vinylaminophosphonate. The product of dehydrophos proteolysis, methyl acetylphosphonate, is thought to be a structural mimic of pyruvate and a possible inhibitor of pyruvate dehydrogenase.

The dehydrophos biosynthetic gene cluster has been identified and cloned in the laboratory of Prof. William W. Metcalf (Department of Microbiology, UIUC) using the phosphonate discovery method outlined in Chapter 1.[7] Briefly, a fosmid library of *S. luridus* was screened for the gene encoding phosphoenolpyruvate (PEP) mutase catalyzing the first enzymatic step in most known phosphonate biosynthetic pathways (Figure 1.3). After sequencing the PEP mutase-positive clones and subsequent gene deletion experiments, the minimal contiguous gene cluster responsible for dehydrophos production was identified. The fosmid containing the putative dehydrophos gene cluster was integrated into *S. lividans*, which allowed for heterologous production of dehydrophos.[7] Bioinformatics analysis identified 16 open reading frames (ORFs), *dhpA-dhpP* (Figure 3.2) within the gene cluster for which protein functions could be proposed by sequence homology to enzymes of known functions. Based on this analysis, a hypothetical biosynthetic pathway for dehydrophos has been proposed (Figure 3.3).[7]

The first two steps resemble the transformations found in other known phosphonate biosynthetic pathways. First, PEP is isomerized to phosphonopyruvate (PnPy) by the action of PEP phosphomutase DhpE, a reaction thermodynamically unfavorable, yet when coupled with subsequent DhpF-catalyzed PnPy decarboxylation ultimately results in the formation of phosphonoacetaldehyde (PnAA). Fe-dependent dehydrogenase DhpG converts PnAA to 2-hydroxyethylphosphonate (2-HEP).[10] As such, the first three steps of the dehydrophos pathway are similar to those found in the biosynthetic pathways of PTT and fosfomycin.[10,11] The early part of the remainder of the proposed biosynthetic steps as well as the later methyl transfer step of the dehydrophos pathway have been validated by analysis of intermediates accumulated by gene-blocked

mutants or by *in vitro* reconstitution of the enzyme activities. This chapter describes elucidation of the remaining steps of dehydrophos biosynthesis. The gene knockout mutants were prepared and their metabolite production was analyzed by Benjamin T. Circello from the Metcalf laboratory. In order to identify the produced metabolites, synthetic standards were prepared in my work and employed as described below. Biochemical approaches were utilized where *in vitro* enzymatic activity provided the substrate scope of the reactions using synthetic substrates and products and their analogues. Most of the phosphonate compounds required for the identification of unknown compounds in NMR or mass spectroscopic (MS) analysis as biosynthetic intermediates in genetic approaches, or substrate and product analogues in biochemical approaches as well as for their use as standards for quantification in MS were not commercially available and were prepared in this work by chemical synthesis. Based on results obtained from these two different approaches, which are often complementary to one another, each step in the proposed hypothetical biosynthetic pathway were confirmed or strongly supported.

Understanding the biosynthesis of phosphonates such as dehydrophos is important in learning the strategies Nature employs for production of novel phosphonates. As well, the biochemistry behind chemically unprecedented transformations of dehydrophos can provide valuable information. For instance, how the two important steps, namely installation of the vinyl group and the tripeptide formation, are achieved in dehydrophos biosynthesis is not clear at this stage. Neither a dehydratase enzyme, which often takes part in vinyl moiety formation, nor peptide-forming enzymes such as nonribosomal peptide synthetases or ATP-grasp enzymes are present in the gene cluster. Therefore,

alternative methods are likely used for these transformations in the dehydrophos-producing strain. Uncovering theses unknown reactions will provide valuable insight in how these transformations, seemingly unique to dehydrophos, are taking place. Furthermore, learning what strategy Nature has adopted to produce novel phosphonates enables further downstream engineering of these biologically active compounds to contain better pharmacokinetic properties towards medicinal applications.

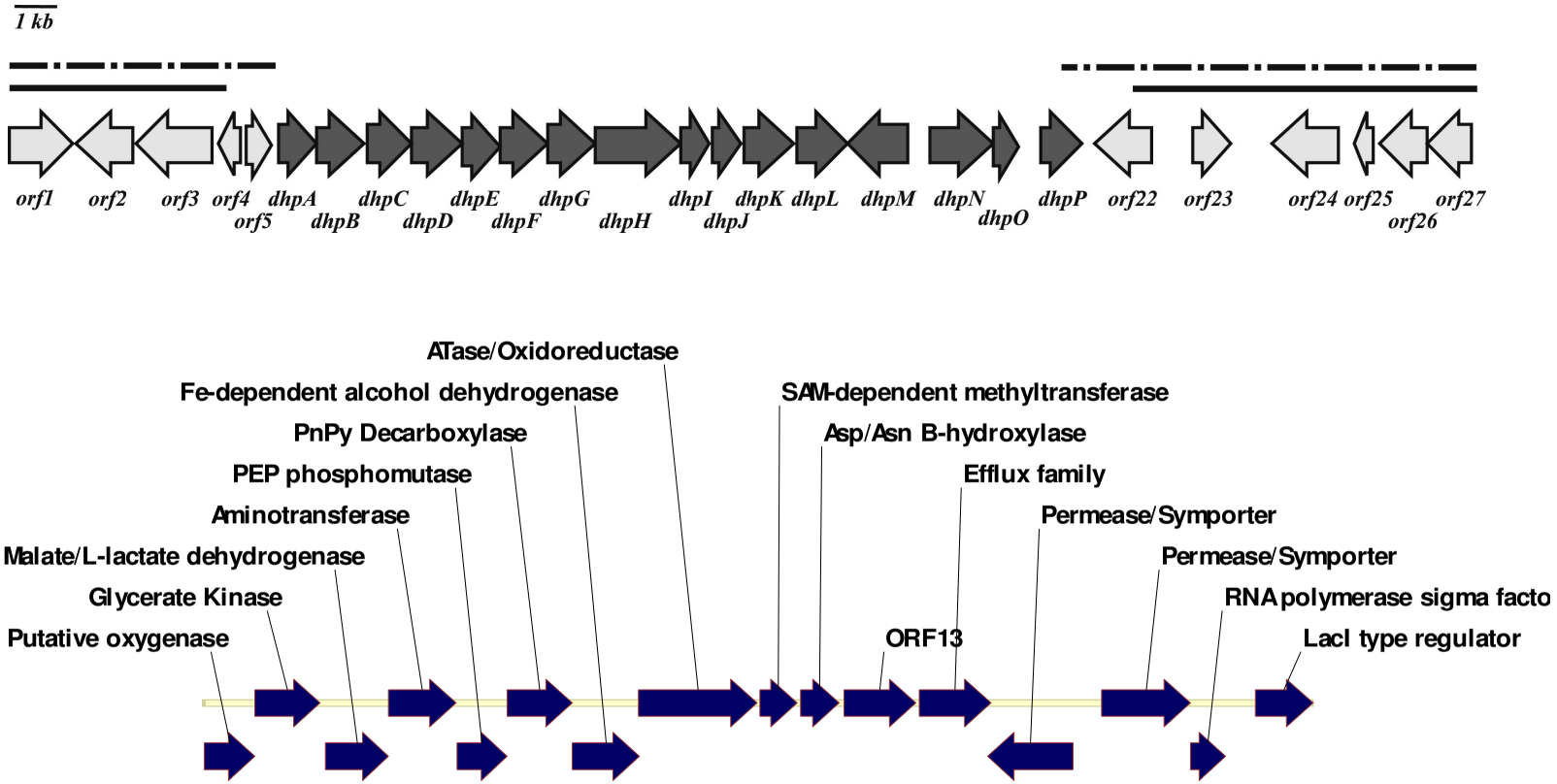


Figure 3.2 The organization of the dehydrophos gene cluster and surrounding DNA. In the first panel, dark arrows indicate ORFs thought to be involved in dehydrophos biosynthesis. The solid lines indicate DNA fragments whose deletion has no effect on production of dehydrophos while the dotted lines indicate deletions that abolish antibiotic production. The second panel shows expansion of the minimal gene cluster, *dhpA* - *dhpP*, with corresponding enzymatic functions assigned using NCBI BLAST analysis. These figures were prepared by Benjamin T. Circello, and have been reproduced with his permission.

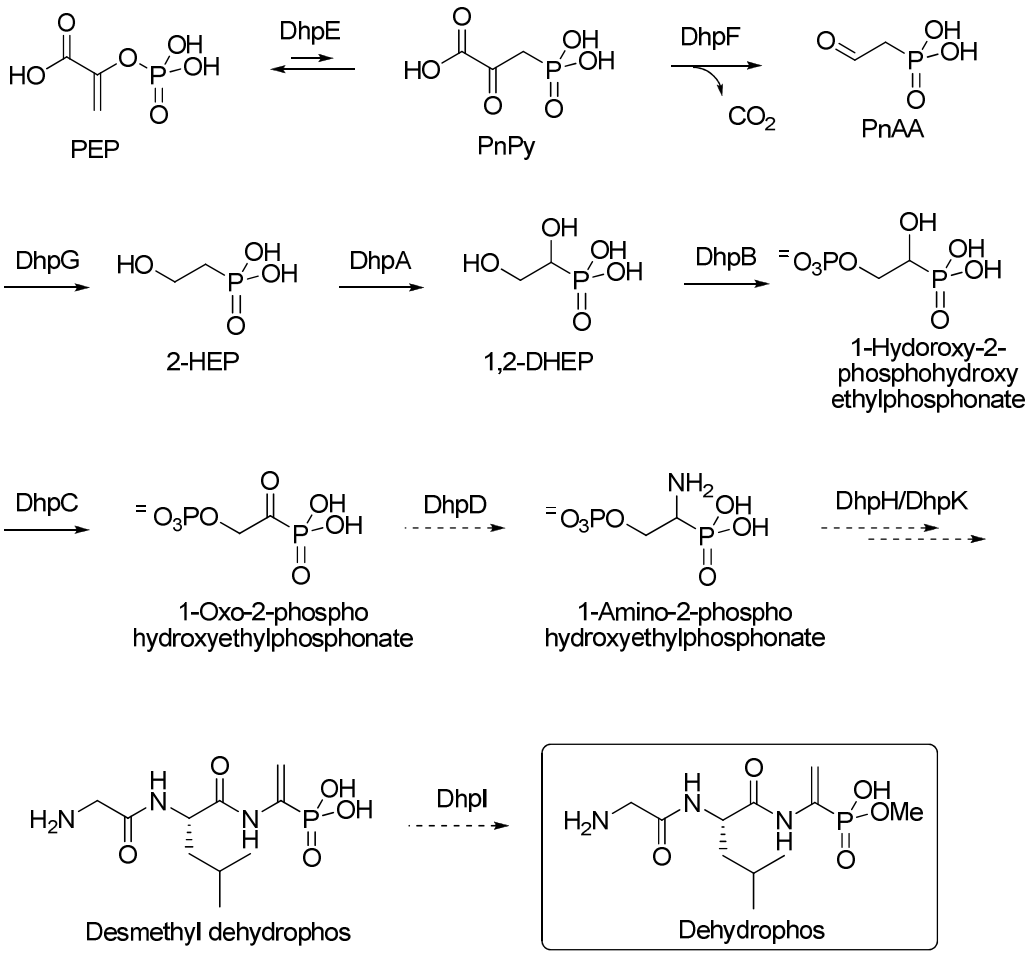


Figure 3.3 Proposed biosynthetic pathway of dehydrophos. Biosynthetic steps that have been confirmed in this work either genetically[7] or biochemically are represented in solid arrows and proposed steps are shown in dashed arrows.

3.2 – Results

3.2.1. Conversion of 2-HEP to 1,2-DHEP by DhpA

Based on the amino acid sequence homology to dioxygenases, DhpA was proposed to catalyze the conversion of the known intermediate 2-HEP to 1,2-dihydroxyethylphosphonate (1,2-DHEP) (Figure 3.3). In order to verify the conversion by DhpA, a phosphonate compound, which accumulated in the spent medium of a *dhpA*-blocked mutant (ΔDhpA) generated by B. Circello, was analyzed by ^{31}P NMR spectroscopy.[7]

^{31}P NMR spectroscopy is a valuable tool in identification of phosphonate compounds whose chemical shifts are within a specific range of 5-30 ppm, well distant from those of the more predominant phosphates and phosphate esters that usually appear in a region more upfield (-20 to 5 ppm). It should be noted, however, that care should be taken when analyzing phosphonate compounds as the phosphorus chemical shifts can vary substantially (up to 5 ppm difference observed) depending on several factors such as pH and ionic strength of the sample. To conclusively verify the identity of a phosphonate, the sample of the unknown phosphonate needs to be mixed with an authentic standard and analyzed under the same ^{31}P NMR condition. A single peak is indicative of the unknown compound being the same molecule as the authentic sample supplemented, whereas appearance of a new peak suggests they are two different molecular species, provided there is no degradation of the phosphonate subjected to the NMR analysis.

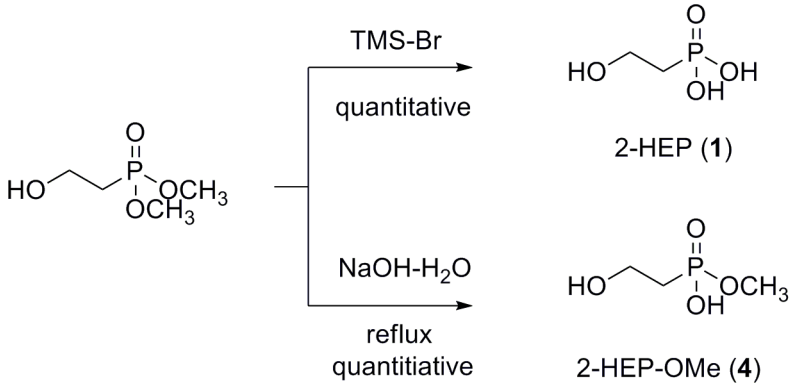
The intermediate 2-HEP (Figure 3.3) was prepared from commercially available dimethyl 2-HEP ester (**1**) by facile bromotrimethylsilane-catalyzed hydrolysis in quantitative yield (Scheme 3.1). The ^{31}P NMR spectrum of 2-HEP displayed a peak with

a chemical shift of 19 ppm, similar to that of the unknown compound that accumulated by the ΔDhpA mutant. When synthetic 2-HEP was added to the spent media of the ΔDhpA mutant, no new phosphonate peak was observed in the ^{31}P NMR spectrum. Therefore, it was confirmed that the unknown phosphonate produced by ΔDhpA is indeed 2-HEP.

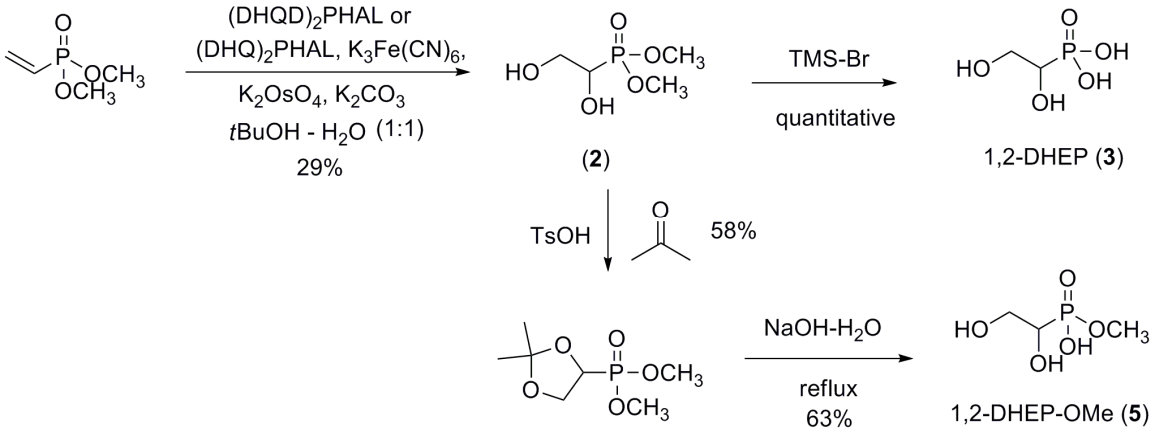
DhpA activity was also analyzed in an *in vitro* enzymatic assay by B. Circello using heterologously produced protein. When 2-HEP was incubated with DhpA in a buffered solution containing α-ketoglutarate, formation of a new phosphonate with a chemical shift at 16 ppm in the ^{31}P NMR spectrum was observed. The dimethyl ester of the anticipated product and putative pathway intermediate 1,2-DHEP was prepared by Sharpless asymmetric dihydroxylation of vinyl phosphonate dimethyl ester utilizing osmium tetraoxide and chiral ligands $(DHQD)_2PHAL$ and $(DHQ)_2PHAL$ (Scheme 3.2). Dimethyl ester (**2**) produced in this manner was subsequently hydrolyzed to the corresponding acid (**3**) using bromotrimethylsilane in a quantitative yield (Scheme 3.2). Enantiomerically enriched 1,2-DHEP was prepared with the purpose of utilization of 1,2-DHEP as a putative substrate in the subsequent enzymatic reactions, where the chirality of the substrate is important. However, since ^{31}P NMR spectroscopy does not differentiate between the two enantiomers it was not necessary to determine the enantiomeric excess of the synthetic product for the experiments described in this chapter. Synthetic product was used in ^{31}P NMR experiments as described above to confirm that the product of the DhpA reaction was indeed 1,2-DHEP.

The possibility of 2-HEP being the substrate for the methyltransferase to be transformed into 2-HEP-OMe prior to the oxygenation to 1,2-DHEP-OMe could not be

ruled out. Therefore, synthesis of putative substrate and product in this alternative scenario, 2-HEP-OMe (**4**) and 1,2-DHEP-OMe (**5**) respectively, were carried out. In order to remove just one of the methyl ester groups, the corresponding dimethyl ester was refluxed in 10% aqueous sodium hydroxide solution until the starting material was completely consumed. This reaction resulted in a quantitative conversion of the dimethyl phosphonate ester to monomethyl phosphonate ester in a facile manner (Scheme 3.2). In case of monodeprotection of dimethyl 1,2-DHEP ester (**2**) to 1,2-DHEP-OMe (**5**), the dihydroxyl group was first masked as an acetonide using acetone in the presence of *p*-toluenesulfonic acid to efficiently isolate the otherwise water-soluble product from the aqueous solvent. Refluxing the acetonide in 10% aqueous sodium hydroxide solution afforded deprotection of the acetonide and cleavage of one of the methyl groups in the dimethyl ester.



Scheme 3.1 Synthesis of 2-HEP (**1**) and 2-HEP-OMe (**4**).



Scheme 3.2 Synthesis of 1,2-DHEP (**3**) and 1,2-DHEP-OMe (**5**).

3.2.2. Conversion of 1,2-DHEP to1,2-DHEP-OPO_3^{2-} by DhpB

Initially, it was speculated that the formation of 1,2-DHEP was followed by its oxidation to 1-oxo-2-hydroxyethyl phosphonate (1-oxo-2-HEP) (Figure 3.4) catalyzed by DhpC (a homolog of malate/lactate dehydrogenases). Such transformation would be very similar to the malate-oxaloacetate interconversion catalyzed by malate dehydrogenase. As such, the gene deletion mutant for ΔDhpC was expected to accumulate a putative DhpC substrate, 1,2-DHEP. The ΔDhpC mutant generated by B. Circello produced two phosphonate intermediates with ^{31}P NMR peaks at approximately 15 ppm, one of which was shown to be 1,2-DHEP using the synthetic standard. Accompanied by these two peaks were the peaks resulted from the upstream intermediates 2-HEP and 2-AEP and phosphate groups. The presence of another phosphonate in this experiment suggests that DhpC may not catalyze the immediate step following DhpA, but another reaction or reactions may occur prior to the DhpC-catalyzed step in the biosynthesis of dehydrophos. Unfortunately, attempts to prepare 1-oxo-2-HEP via synthetic methods described below (Scheme 3.3) were not successful to provide clear evidence of the presence of 1-oxo-2-HEP in the DhpC catalyzed reaction utilizing 1,2-DHEP as substrate. However, when DhpC culture supernatant was treated with alkaline phosphatases in B. Circello's experiment, 1,2-DHEP resulted suggesting the other phosphonate produced by ΔDhpC is 1,2-DHEP phosphate ester.

In my attempts to obtain 1-oxo-2-HEP, glycolic acid was protected with *tert*-butyldimethylsilyl (TBDMS) chloride in the presence of imidazole followed by conversion to the acyl chloride by oxalyl chloride. This chloride was subjected to Arbuzov-type addition of phosphite[12] but the desired product could not be obtained.

Similarly, benzyloxyacetyl chloride was directly subjected to an Arbuzov-type phosphite addition reaction followed by the cleavage of the benzyl ether and methyl esters but the desired product could not be obtained. According to the ^{1}H NMR a peak with a large coupling constant was observed suggesting the presence of a hydrogen phosphonate species in the product mixture.

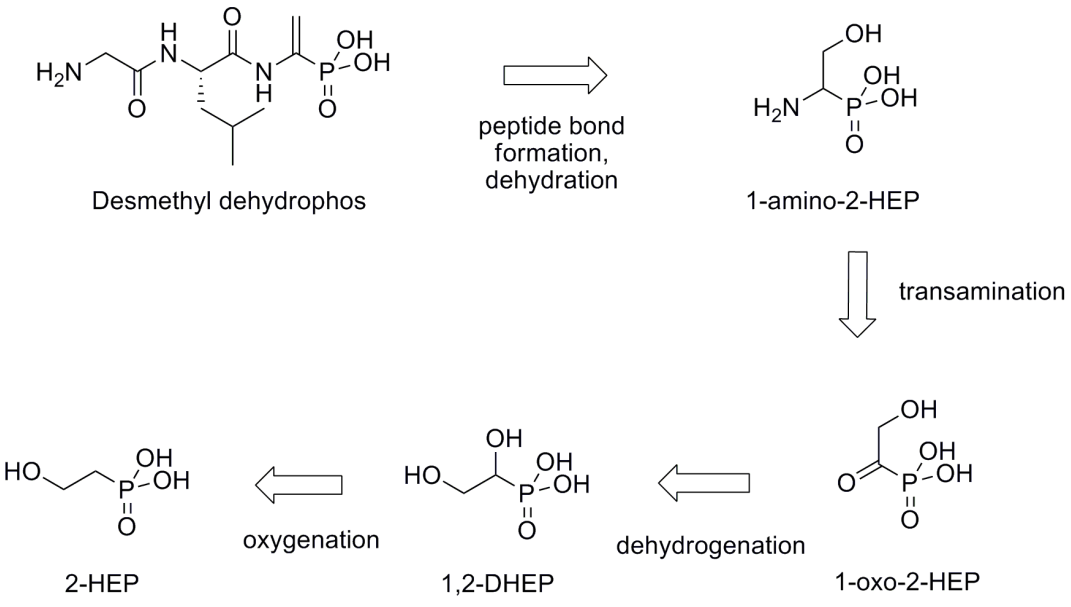
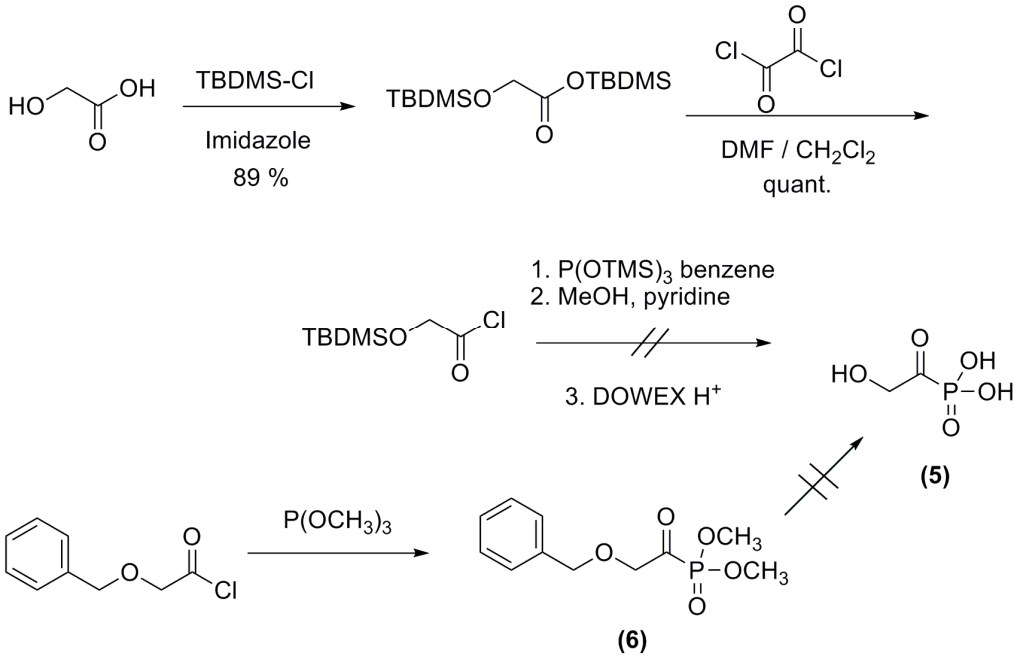


Figure 3.4 Retrosynthesis outlining series of enzymatic reactions leading to desmethyl dehydrophos starting from 2-HEP.



Scheme 3.3 Attempted synthetic routes to 1-oxo-2-HEP.

The protein product of *dhpB* is homologous to glycerate kinases, and it can be perceived that this enzyme takes 1,2-DHEP as a substrate to phosphorylate the hydroxyl group at the C2 position. The ΔDhpB mutant indeed produced 1,2-DHEP in a gene-blocked mutant experiment performed by B. Circello.[7] My attempt to express DhpB with an *N*-terminal hexahistidine tag from a heterologous host was not successful resulting in poor expression. Moreover, the small amounts of DhpB overexpressed (confirmed by Western Blot with anti-His_6) at an optimized induction condition produced insoluble protein. Coexpression with chaperone proteins such as DnaK, DnaJ and GrpE using a commercially available chaperone plasmid, pG-KJE8 or employing the betaine-sorbitol osmotic stress system, which often improves protein solubility, did not improve the solubility of DhpB. Finally, a DhpB construct in which the protein was fused to the maltose binding protein (MBP) at its *N*-terminus was prepared, but soluble protein could not be obtained. Alternative attempts towards *in vitro* reconstitution of DhpB, which may support the hypothesis of 1,2-DHEP being the physiological substrate and its phosphate ester the product, are still under investigation.

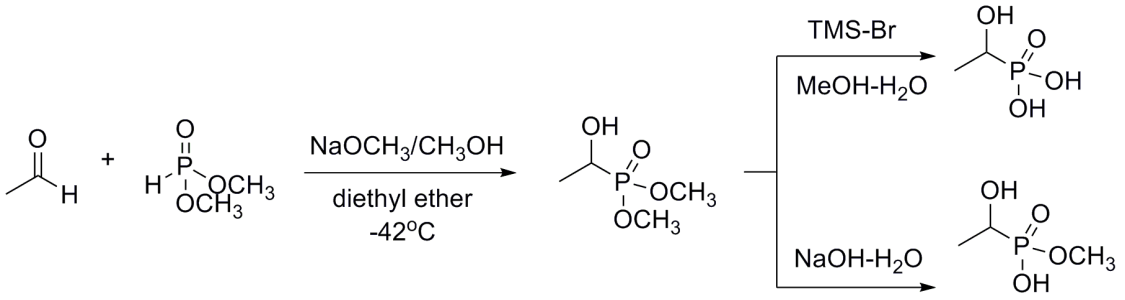
3.2.3. Production of 1-HEP-OMe, an Unexpected Metabolite

During the gene-knockout experiments carried out by B. Circello, an unknown compound with a peak at 24.5 ppm by ^{31}P NMR was observed from the growth media of two different strains. One was a mutant strain ΔDhpK, and the other one was a lacI repressor-deleted mutant strain. Synthetically available phosphonate intermediates proposed to form during the biosynthesis were used as authentic standards in the ^{31}P NMR method described previously, but none of these compounds matched the unknown at 24.5 ppm. The unknown phosphonate compound in the spent growth media could also not be isolated efficiently for further analysis. Instead, I carried out a 2-dimensional ^{1}H-^{31}P Heteronuclear Multiple Bond Correlation (HMBC) NMR experiment to obtain information on the types of protons that are near to phosphorus.

^{1}H-^{31}P HMBC NMR of the intermediate is shown in Figure 3.5. which was used to identify the chemical shifts of protons coupled to the ^{31}P nucleus (i.e. less than 4 bonds away). This technique is selective and relieves overlap by proton peaks arising from solvent, nutrients, and other non-phosphonate metabolites. Only three proton peaks which were coupled to the phosphorus peak at 24.5 ppm were observed (Figure 3.5A). In the far upfield region, a doublet of a medium height at 1.3 ppm was observed in addition to a strong peak at 3.6 ppm and a small peak at 3.9 ppm. The strength of the signal can be a result of the combination of the number of bonds separating the nuclei as well as the number of protons contributing to the peak. Using the known chemical shifts of protons in the ethylphosphonate backbone with putative functional groups installed at various positions, several chemical structures including 1-HEP-OMe were proposed as a structure

of the unknown compound. 1-HEP-OMe would also explain the complex ^{31}P NMR spectrum.

In order to verify this hypothesis, 1-HEP-OMe was synthesized (Scheme 3.4). Condensation of acetaldehyde and dimethyl phosphite produced 1-HEP dimethyl ester, which was subsequently treated with iodotrimethylsilane or sodium hydroxide solution to afford 1-HEP and 1-HEP-OMe, respectively, both in quantitative yield. The ^{1}H-^{31}P HMBC NMR for 1-HEP-OMe was taken (Figure 3.5B). The spectrum matched well with the spectrum of the sample of the intermediate produced by the ΔDhpK mutant (Figure 3.5A), and when the two samples were mixed and analyzed by ^{31}P NMR spectroscopy, only one phosphorus peak at 24.5 ppm could be observed showing the identity of the unknown to be 1-HEP-OMe. Based on the observation that dehydrophos readily breaks down into acetyl phosphonate monomethyl ester and a dipeptide by cellular peptidase hydrolysis, it was proposed that 1-HEP-OMe is a metabolite generated by an endogenous lactate dehydrogenase activity on acetyl phosphonate monomethyl ester (Figure 3.5). Using acetyl phosphonate monomethyl ester and 1-HEP-OMe which were synthesized by myself according to the literature procedure,[8] Ben Circello was able to show that the dehydrophos indeed generated 1-HEP-OMe after incubation in *S. lividans* cell culture media.



Scheme 3.4 Synthesis of 1-HEP and 1-HEP-OMe.

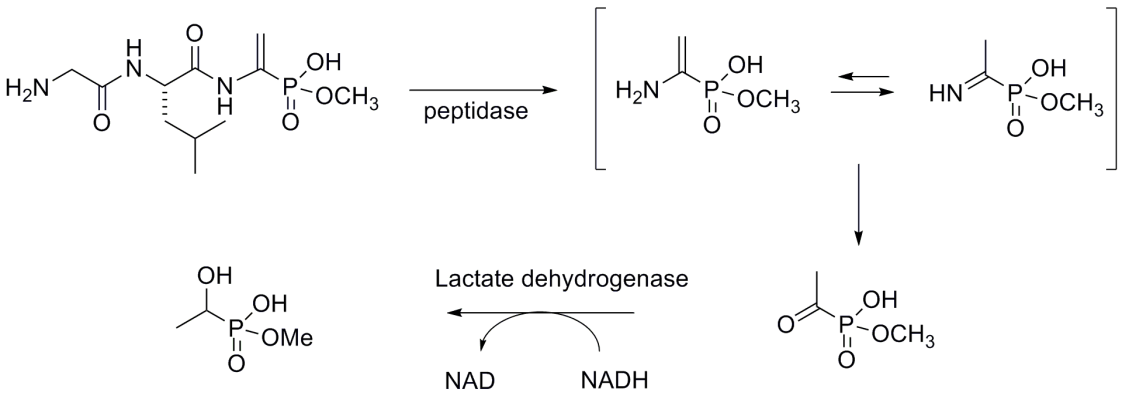


Figure 3.5 A proposed route to the formation of 1-HEP-OMe.

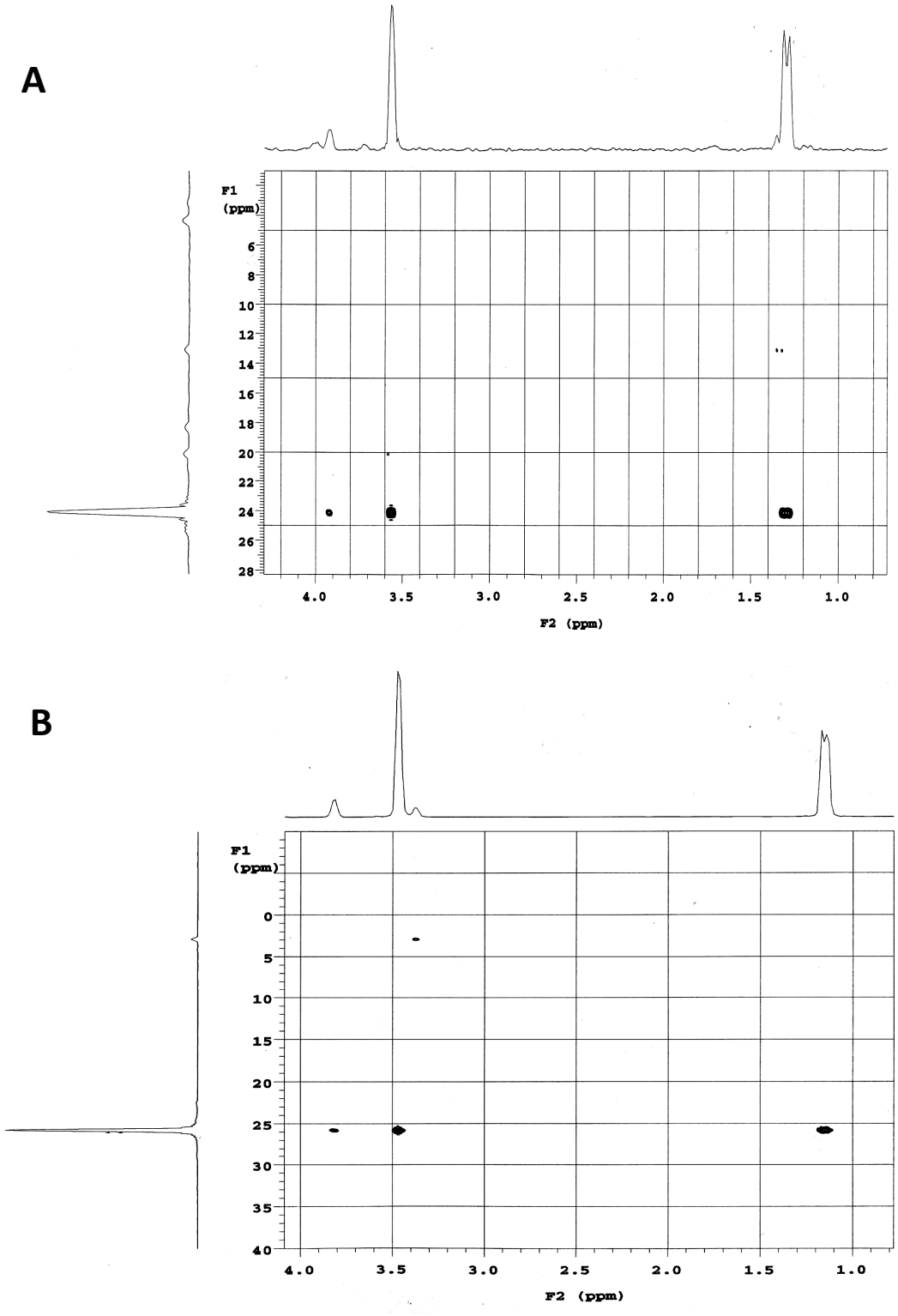


Figure 3.6 2-Dimensional ^{1}H-^{31}P HMBC NMR. (A) A spectrum taken from a sample of concentrated spent medium in which ΔDhpK was grown and (B) a spectrum of an authentic synthetic 1-HEP-OMe sample.

3.3 – Discussion and Conclusion

As part of a multidisciplinary project investigating naturally occurring phosphonates with biological activity, this chapter describes a successful application of using the PEP mutase gene as a phosphonate marker for identifying and isolating a phosphonate gene cluster. Both genetic and biochemical approaches were applied in the elucidation of the dehydrophos biosynthetic pathway outlined in this chapter. The early steps in the dehydrophos biosynthesis are shared by biosynthetic steps of other phosphonates including fosfomycin and PTT. Due to the high homology of DhpE and DhpF to the corresponding PEP mutase and PnPy decarboxylase involved in the first two transformations of most known phosphonate compounds, their roles were assigned to be PEP mutase-catalyzed conversion of PEP to PnPy and the subsequent decarboxylation to PnAA. Similarly DhpG was assigned a function of Fe-dependent alcohol dehydrogenase as often observed in other phosphonate biosynthesis, which catalyze conversion of PnAA to 2-HEP.[10]

Based on the structure of dehydrophos, protein functions predicted from their amino acid sequence using NCBI's BLAST search, and the chemical logic starting from 2-HEP, a biosynthetic pathway for dehydrophos was proposed. Using this proposed pathway and putative substrate and product in each step, genetic and biochemical analysis was performed. In the genetic approaches conducted by B. Circello, phosphonate compounds produced by gene-blocked mutants were analyzed using ^{31}P NMR without purification, and their identity was determined by comparison with authentic synthetic compounds available via synthesis. When the compounds could not be identified in this

manner, ^{1}H-^{31}P HMBC NMR was used to obtain information on the chemical structure of the unknown compound.

Several phosphonate compounds such as 2-HEP, 2-HEP-OMe, 1,2-HEP, 1,2-HEP-OMe, 1-HEP, 1-HEP-OMe, acetyl phosphonate, and acetyl phosphonate-OMe used in this study were prepared successfully by chemical synthesis. Not only have these compounds served as synthetic standards in elucidating the dehydrophos biosynthetic pathway, but these compounds also make up a phosphonate library with applications in biosynthetic studies of other important phosphonates. This study demonstrates that DhpA converts 2-HEP to 1,2-DHEP by DhpA-catalyzed oxygenation as supported by both gene deletion and reconstitution of heterologously expressed DhpA using 2-HEP as substrate in the presence of α-ketoglutarate. Although activity of DhpB could not be reconstituted *in vitro*, the gene deletion experiment of DhpB, which resulted in accumulation of 1,2-DHEP, as well as the gene deletion study of DhpC, which showed production of a phosphate ester of 1,2-DHEP, suggests the step following DhpA is DhpB-catalyzed phosphorylation of the 2-hydroxyl group in a similar reaction to glycerate kinases. While biochemical evidence is still to be obtained, the next step is suspected to be catalyzed by malate dehydrogenase, DhpC, which may convert 1,2-DHEP-OPO_3^{2-} to 1-oxo-2-HEP-OPO_3^{2-}.

While the remaining steps are yet to be elucidated, this study was able to provide a synthetic, biochemical, and genetic approach to discovery of novel enzymatic transformations as well as new phosphonate intermediates. Having the knowledge of phosphonate biochemistry in hand, further engineering of these unidentified

phosphonates may provide a reservoir for new compounds with useful biological activities.

3.4 – Experimental Procedures

3.4.1. Chemical Syntheses of Phosphonates

General Materials and Methods for Chemical Synthesis. All NMR spectra were recorded on a Varian Unity 400 or a Varian Unity 500 spectrometer. ^{1}H NMR spectra were referenced to TMS at 0 ppm or H_2O at 4.67 ppm, and ^{13}C NMR spectra were referenced to $CHCl_3$ at 77.0 ppm. For ^{31}P NMR spectra 85% phosphoric acid in D_2O was used as an external reference (0 ppm). Fractions collected during silica gel column chromatography were analyzed using thin layer chromatography (TLC). Unless otherwise specified, all compounds and solvents were obtained from Fisher Scientific or Sigma-Aldrich. CH_2Cl_2 was distilled over CaH_2 immediately prior to use.

Dehydrophos. Dehydrophos was synthesized and purified according to a literature procedure[3] and utilized within a month after its preparation due to its decomposition overtime.

2-Hydroxyethylphosphonate methyl ester (**4**). To an aqueous solution of 10% sodium hydroxide (10 mL) was added dimethyl 2-hydroxyethylphosphonate dimethyl ester (**4**) (6.49 mmol), and the reaction was refluxed for 1 h. After cooling to room temperature, the reaction mixture was acidified with 2 M HCl, concentrated, and redissolved in 20 mL of methanol. The suspension was passed through a plug of celite to afford the desired

product in a quantitative yield. ^{1}H NMR (400 MHz, $CDCl_3$) δ 3.70 (2H, dt, *J* = 14.4, 7.2 Hz, C*H*$_2$OH), 3.35 (3H, d, *J* = 11.2 Hz, OC*H*$_3$), 2.01 (2H, dt, *J* = 18.4, 7.2 Hz, PC*H*$_2$); ^{13}C NMR (125.6 MHz, D_2O) δ 56.4 (HO*C*H$_2$), 51.8 (O*C*H$_3$, d, *J* = 5.5 Hz), 28.6 (*C*H$_2$P, d, *J* = 132.5 Hz); ^{31}P NMR (202.3 MHz) δ 28.6. HRMS (EI) calcd for $C_3H_9O_4P$ 140.0239 found 140.0228.

1,2-Dihydroxyethylphosphonate dimethyl ester (**2**).[13] To a suspension of AD mix *β** (6.02 g) in 50 mL of 1:1 *tert*-butanol/H_2O, which had been stirred for 15 min at room temperature, was added dropwise dimethyl vinylphosphonate (4.3 mmol, Alfa Aesar) and the solution was stirred for 4 days at room temperature. The yellow suspension was placed in an ice bath and quenched by the addition of sodium sulfite (5.0 g) upon which the solution turned green. After 20 min the ice bath was removed, and the suspension was stirred at room temperature for another 20 min. The solution was extracted with ethyl acetate (3 × 50 mL), and the combined organic layer was dried over magnesium sulfate. After evaporating the excess solvents and reagents, the crude product was purified by silica gel column chromatography (EtOAc:MeOH = 4:1, R_f=0.4) to afford the desired product in 29% yield as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.13-4.10 (1H, m, PC*H*), 3.91-3.85 (2H, m, C*H*$_2$), 3.83 (3H, d, *J* = 10.5 Hz, OC*H*$_3$), 3.82 (3H, d, *J* = 10.5 Hz, OC*H*$_3$); ^{13}C NMR (125.6 MHz, $CDCl_3$) δ 68.8 (*C*H, d, *J* = 160 Hz), 62.3, 53.3 (O*C*H$_3$, d, *J* = 24.9 Hz); ^{31}P NMR (202.3 MHz, $CDCl_3$) δ 26.9. (Stereospecific dihydroxylation by AD mix *β* yields enantiomerically enriched product, however, NMR experiments described in this report do not differentiate between enantiomers, and

therefore it was not necessary for this report to determine the enantiomeric excess of the product.)

1,2-Dihydroxyethylphosphonate monomethyl ester (**5**). To a 10 mL round-bottomed flask containing the ester **4** (167 mg, 0.982 mmol) was added 4.5 mL of acetone and the reaction was stirred at room temperature. To this suspension was added *p*-toluene sulfonic acid (9.3 mg, 5 mol%) and the mixture was stirred overnight. The temperature was raised on an oil bath, and the reaction mixture was refluxed for 2 h. After cooling down, the crude products were extracted with dichloromethane (3 × 40 mL), and the combined organic layer was dried over magnesium sulfate. Purification was achieved by silica gel column chromatography (EtOAc:MeOH = 10:1) to isolate a colorless oil. To this material was added 10% NaOH solution and the reaction was heated to 100 ^{o}C. After refluxing for 3 h, the reaction was acidified using 1 M HCl, and passed through a Dowex (H^+) column to yield the desired acid in a 36.1% overall yield. ^{1}H NMR (500 MHz, D_2O) δ 3.75 (1H, td, *J* = 9.5, 2.5 Hz, PC*H*OH), 3.68 (1H, ddd, *J* = 12, 5.5, 2.5 Hz, CH_2OH), 3.50 (1H, ddd, *J* = 12, 9.5, 4.5 Hz, CH_2OH), 3.45 (3H, d, *J* = 10.5 Hz, OCH_3); ^{13}C NMR (125.6 MHz, D_2O) δ 68.58 (d, *J* = 157 Hz, HO*C*HP), 61.84 (d, *J* = 9.2 Hz, PO*C*H_3), 52.61 (d, *J* = 6.5 Hz, HO*C*H_2); ^{31}P NMR (202.3 MHz, D_2O) δ 20.8. HRMS (ESI) calcd for $C_3H_9O_5NaP$ $[M + Na]^+$ 179.0085 found 179.0078.

1,2-Dihydroxyethylphosphonic acid (**3**). To a 50 mL round-bottomed flask were added 91.8 mg (0.54 mmol) of the ester **7** and 5 mL of dry dichloromethane. While stirring at room temperature, bromotrimethylsilane (380 mg, 3.5 equiv) was added dropwise. After

stirring further at room temperature for 2 h, 2 mL of methanolic water was added. The excess solvents and reagents were removed under reduced pressure and after diluting with water, the product was mixed with charcoal and filtered through celite to obtain the desired product in a quantitative yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.68 (1H, ddd, *J* = 10.0, 9.0, 3.0 Hz, PC*H*), 3.60 (1H, ddd, *J* = 12.0, 7.0, 3.0 Hz, CH_2), 3.43 (1H, ddd, *J* = 12.0, 9.0, 5.5 Hz, CH_2); ^{13}C NMR (125.6 MHz, D_2O) δ 69.1 (HO*C*P, d, *J* = 156.2 Hz), 61.9 (HO*C*H_2, d, *J* = 8.8 Hz); ^{31}P NMR (202.3 MHz, D_2O) δ 21.3.

1-Hydroxyethylphosphonate dimethyl ester. A solution of acetaldehyde (14.31 mmol) and dimethyl phosphite, 20 mmol) in dry diethyl ether (20 mL) was cooled to -42 ^{o}C, and a saturated solution of sodium methoxide (0.14 mmol) in methanol was added dropwise. The reaction mixture was stirred for 10 min, and the excess solvent and reagent were removed at reduced pressure. The crude products were extracted with EtOAc (3 × 30 mL), and the organic layer was dried over $MgSO_4$. After filtering off the drying agent, solvent was removed on the rotovap to afford the desired product in near quantitative yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.08 (1H, dq, *J* = 7.0, 3.0 Hz, HOC*H*), 3.82 (6H, dd, *J* = 10.5, 4.5 Hz, POCH_3), 1.45 (3H, dd, *J* = 17.5, 7 Hz, CH_3CHOH); ^{31}P NMR (202.3 MHz, $CDCl_3$) δ 29.0.

1-Hydroxyethylphosphonate monomethyl ester. To a solution of dimethyl 1-hydroxyethyl phosphonate ester (1.655 mmol), was added 10% NaOH dropwise while stirring. The reaction mixture was refluxed for 4.5 h. After cooling, the crude product was acidified with 1 M HCl, and extracted with chloroform (3 × 40 mL), and dried over

sodium sulfate. After the solvent was removed by reduced pressure, the desired product was obtained in near quantitative yield. . ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.85 (1H, m, HOC*H*), 3.52 (3H, d, *J* = 10.5 Hz, POCH_3), 1.17 (3H, dd, *J* = 9.5, 7.5, Hz, CH_3CHOH); ^{13}C NMR (125.6 MHz, D_2O) δ 63.2 (*C*HOH, d, *J* = 161.3 Hz), 52.5 (PO*C*H_3, d, *J* = 6.9 Hz), 16.8 (*C*H_3); ^{31}P NMR (202.3 MHz, $CDCl_3$) δ 27.0. HRMS (ESI) calcd for $C_3H_9O_4NaP$ $[M + Na]^+$ 163. 0136 found 163.0141.

3.5 – References

1. Johnson, R. D.; Kastner, R. M.; Larsen, S. H.; Ose, E. E.: *Antibiotic A53868 and process for production thereof*. US Patent **1984**, *4*, 482,488.

2. Hunt, A. H.; Elzey, T. K.: "Revised structure of A53868a." *J. Antibiot.* **1988**, *41*, 802-802.

3. Whitteck, J. T.; Ni, W.; Griffin, B. M.; Eliot, A. C.; Thomas, P. M.; Kelleher, N. L.; Metcalf, W. W.; van der Donk, W. A.: "Reassignment of the structure of the antibiotic A53868 reveals an unusual amino dehydrophosphonic acid." *Angew. Chem. Int. Ed.* **2007**, *46*, 9089-9092.

4. Metcalf, W. W.; van der Donk, W. A.: "Biosynthesis of phosphonic and phosphinic acid natural products." *Annu. Rev. Biochem.* **2009**, *78*, 65-94.

5. Kugler, M.; Loeffler, W.; Rapp, C.; Kern, A.; Jung, G.: "Rhizocticin A, an antifungal phosphono-oligopeptide of *Bacillus subtilis* ATCC6633: biological properties." *Arch. Microbiol.* **1990**, *153*, 276-281.

6. Atherton, F. R.; Hall, M. J.; Hassall, C. H.; Lambert, R. W.; Lloyd, W. J.; Lord, A. V.; Ringrose, P. S.; Westmacott, D.: "Phosphonopeptides as substrates for peptide-transport systems and peptidases of *Escherichia coli*." *Antimicrob. Agents Chemother.* **1983**, *24*, 552-528.

7. Circello, B. T.; Eliot, A. C.; Lee, J.-H.; van der Donk, W. A.; Metcalf, W. W.: "Molecular cloning and heterologous expression of the dehydrophos biosynthetic gene cluster." *Chem. Biol.* **2010**, *17*, 402-411.

8. Obrien, T. A.; Kluger, R.; Pike, D. C.; Gennis, R. B.: "Phosphonate analogs of pyruvate: Probes of substrate binding to pyruvate oxidase and other thiamin pyrophosphate-dependent decarboxylases." *Biochim. Biophys. Acta* **1980**, *613*, 10-17.

9. Ntai, I.; Bachmann, B. O.: "Identification of ACE pharmacophore in the phosphonopeptide metabolite K-26." *Bioorg. Med. Chem. Lett.* **2008**, *18*, 3068-3071.

10. Shao, Z. Y.; Blodgett, J. A. V.; Circello, B. T.; Eliot, A. C.; Woodyer, R.; Li, G. Y.; van der Donk, W. A.; Metcalf, W. W.; Zhao, H. M.: "Biosynthesis of 2-hydroxyethylphosphonate, an unexpected intermediate common to multiple phosphonate biosynthetic pathways." *J. Biol. Chem.* **2008**, *283*, 23161-23168.

11. Blodgett, J. A. V.; Thomas, P. M.; Li, G. Y.; Velasquez, J. E.; van der Donk, W. A.; Kelleher, N. L.; Metcalf, W. W.: "Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide." *Nat. Chem. Biol.* **2007**, *3*, 480-485.

12. Glabe, A. R.; Sturgeon, K. L.; Ghizzoni, S. B.; Musker, W. K.; Takahashi, J. N.: "Novel functionalized acylphosphonates as phosphonoformate analogs." *J. Org. Chem.* **1996**, *61*, 7212-7216.

13. Sharpless, K. B.; Amberg, W.; Bennani, Y. L.; Crispino, G. A.; Hartung, J.; Jeong, K. S.; Kwong, H. L.; Morikawa, K.; Wang, Z. M.; Xu, D. Q., et al.: "The osmium-catalyzed asymmetric dihydroxylation - a new ligand class and a process improvement." *J. Org. Chem.* **1992**, *57*, 2768-2771.

CHAPTER 4: CHARACTERIZATION OF A NOVEL PHOSPHONATE *O*-METHYLTRANSFERASE INVOLVED IN DEHYDROPHOS BIOSYNTHESIS

4.1 – Introduction

Phosphonic acids exhibit a range of biological activities due to their ability to mimic naturally prevalent compounds including phosphate esters (Figure 1.2*A*) and carboxylates (Figure 1.2*B*). Despite their ability to inhibit many enzymatic reactions related to therapeutic applications, phosphonic acids have not been fully explored as drug candidates. A general drawback of phosphonates is their high charge state at neutral pH, and therefore poor bioavailability, which renders them less than ideal for achieving high concentrations at the target site. Two different strategies are found in Nature to overcome these limitations. The first strategy can be found in phosphinothricin which mimics the tetrahedral intermediate generated in glutamine synthetase by using a phosphinate group containing two P-C bonds (Figure 1.2*C*).[1,2] In an alternative solution, in dehydrophos the phosphonate group is esterified, a common approach in synthetic prodrugs[3] but the first such example in a natural product (Figure 3.3).[4]

The structure and biosynthesis of dehydrophos were discussed in the previous chapter of this dissertation. Dehydrophos contains a unique vinyl aminophosphonate moiety linked to a Gly-Leu dipeptide (Figure 3.3). The mode of action of dehydrophos is currently not known, but analogous to other peptide antibiotics, it is anticipated that after uptake by a peptide transporter, a peptidase will release the active phosphonate species.[5,6] The methyl ester of vinyl aminophosphonic acid, the hydrolysis product of the tripeptide, is believed to be responsible for the antibiotic activity of dehydrophos as this enamine

tautomerizes to an imine, and then undergoes spontaneous hydrolysis to form acetylphosphonic acid methyl ester, a potent inhibitor for pyruvate dehydrogenase.[7]

One notable transformation in the biosynthetic pathway of dehydrophos is the unprecedented *O*-methylation of the phosphonic acid that results in the phosphonate monomethyl ester which is characteristic for dehydrophos. This transformation was hypothesized to be carried out by an *S*-adenosyl-L-methionine (SAM or AdoMet) dependent methyltransferase encoded by *dhpI* (Figure 4.1).[8] Amino acid sequence searches using BLAST as well as conserved protein domain searches by InterProScan indicate that DhpI is a Type I methyltransferase utilizing SAM as a methyl donor. Previous structural studies have shown that this protein family forms a typical Rossman-like alpha-beta fold. The ClustalW sequence alignments with proteins of known function show no sequence homology in the putative substrate binding domain, indicating that the substrate of this enzyme is unique to this protein. Although numerous SAM-dependent *O*-methyltransferases are known to methylate various functional groups including carboxylic acid and hydroxyl groups, the methylation of phosphonic acids by this group of enzymes has not been reported. The fact that methyl acetylphosphonate inhibits pyruvate dehydrogenase 125 times more potently than acetylphosphonate[7] may provide a rationale for methylation in dehydrophos as a mean to improve its bioavailability and activity.

In order to examine the substrate scope of DhpI, the phosphonate intermediates in the putative biosynthetic pathway were interrogated as potential substrates for DhpI. At the outset of my studies, the exact timing of the methyl transfer reaction was not clear as any of the intermediates shown in Figure 4.2 could be the natural substrate for this *O*-

methyltransferase. With overexpressed and purified DhpI in hand, its enzymatic activity with various synthetic substrates could be studied. In parallel to the study of the DhpI reaction, several crystal structures of DhpI were produced by our collaborators to further elucidate the determinants of substrate recognition. With structural information, the amino acid residues in proximity to the substrate binding pocket site could be identified, and the roles of those residues were examined by site-directed mutagenesis in my work. Furthermore, the potential in applications of DhpI as a versatile biocatalyst have been demonstrated by showing the broad substrate tolerance in esterifying other biologically active phosphonates, such as the antibiotic fosfomycin and antimalarial agent fosmidomycin.

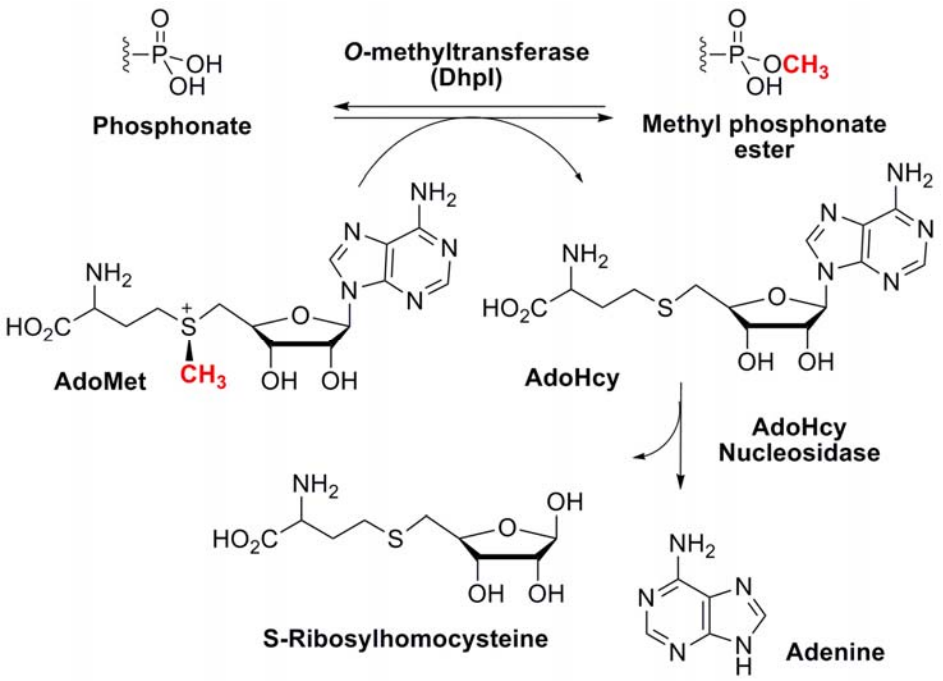


Figure 4.1 Schematic diagram of the SAM-dependent methyl transfer reaction catalyzed by DhpI where a phosphonic acid is converted to a phosphonate methyl ester in the SAH nucleosidase coupled assay. AdoHcy, the by-product of the reaction is a strong product inhibitor.

4.2 – Results

4.2.1. Expression and Purification of Recombinant Proteins for DhpI Assays.

In order to investigate the methyltransferase activity, a recombinant system that allows production of DhpI in a soluble and easily isolatable form was prepared. The *dhpI* gene was cloned from *Streptomyces luridus* genomic DNA and inserted into the expression vector pET15b. DhpI was heterologously expressed in *E. coli* with an *N*-terminal hexahistidine tag (His_6-DhpI). The expression of the 26 kDa protein was achieved in *E. coli* Rosetta 2 (DE3), which is a BL21 DE3 derived strain that can supply tRNAs for rare codons not highly produced by *E. coli*. Successful purification by immobilized nickel affinity chromatography afforded an overall yield of 10-15 mg of purified His_6-DhpI per liter of culture (Figure 4.2).

In vitro reconstitution of DhpI activity was carried out in the presence of a coupling enzyme, SAH nucleosidase. Most of the reported SAM-dependent methyltransferases undergo a strong product inhibition by *S*-adenosyl-L-homocysteine (SAH or AdoHcy, Figure 4.1).[9] This can be circumvented by either using a high concentration of SAM in the assay, or by removing SAH from the reaction mixture as soon as it is made. In this study the latter method was employed by using AdoHcy nucleosidase, which converts SAH to adenine and *S*-ribosylhomocysteine (Figure 4.1). The recombinant *E. coli* SAH nucleosidase was overexpressed as a fusion protein with an *N*-terminal hexahistidine tag in *E. coli* BL21, allowing facile purification by using Ni-affinity column chromatography (Figure 4.2). In order to alleviate product inhibition from AdoHcy, SAH nucleosidase utilized in the assay is required to be highly efficient. The catalytic efficacy of *E. coli* SAH nucleosidase reported in the literature is 11.6×10^6

M^{-1} s^{-1}, and in general the enzyme has much higher activity than that of most known methyltransferases.[9] The efficacy of heterologously expressed and purified SAH nucleosidase was confirmed by LC-MS detection of the disappearance of the SAH peak (385, M+1) as well as by complete loss of the SAH signal in the diode array detector (DAD) coupled to HPLC. The SAH breakdown was observed only when SAH nucleosidase was present in the assay.

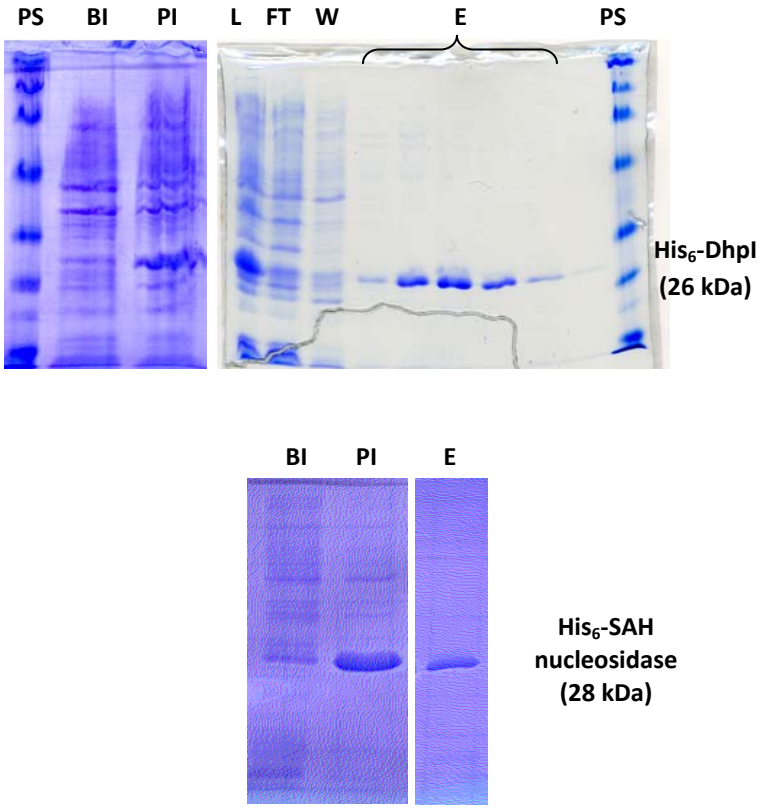


Figure 4.2 Typical overexpression and purification gel for DhpI and SAH nucleosidase. The 10% SDS-PAGE gels were stained with Coomassie Brilliant Blue. The size of the protein was determined by comparing to protein standards (PS = protein standard, BI = pre-induction sample, PI = post-induction sample, L = proteins in the lysed cell, FT = unbound proteins eluting from the Ni^{2+} column, W = wash fractions, E = proteins eluted from the IMAC column with buffer containing a high imidazole concentration).

4.2.2. In Vitro Reconstitution of Enzymatic Activity of DhpI.

The first evaluation of the methylation activity of DhpI focused on the conversion of the early biosynthetic intermediates 2-HEP and 1,2-dihydroxyethylphosphonic acid (DHEP) to their monomethyl phosphonate esters, 2-HEP-OMe and DHEP-OMe. The methylated compound 2-HEP-OMe was previously shown to accumulate in the spent growth medium of some of the genetic mutants generated by B. Circello in the heterologous host *Streptomyces lividans*.[8] 2-HEP and DHEP were chemically synthesized as described in Chapter 3 and the rate of formation of the methyl ester products was determined by direct detection using HPLC coupled to atmospheric pressure chemical ionization mass spectrometry (APCI/MS) with synthetic compounds as authentic standards (Figure 4.3). Additional transformation to dimethyl phosphonate esters was not observed even after a prolonged reaction time. By direct comparison with the authentic sample of the methylated product, both ^{31}P NMR spectroscopy and the fragmentation pattern in tandem MS confirmed that the methylation occurred on the oxygen of the phosphonic acid, and not on the hydroxyl groups of DHEP and 2-HEP.

Although both 2-HEP (Figure 4.3*A*) and DHEP were methylated by DhpI, they proved relatively poor substrates that were unable to saturate the enzyme even at high concentrations. *N*-Acetyl-1-aminoethylphosphonic acid (Ac-1-AEP or AcAlaP) was also converted to the corresponding methyl ester whereas little or no conversion was observed for 1-AEP and 2-AEP. Surprisingly, 1-hydroxyethylphosphonic acid (1-HEP) was a much better substrate (Table 4.1), and as discussed in Chapter 3 its methyl ester (1-HEP-OMe) was also produced by some of the mutants of the producing strain.[8] However, it is difficult to envision a biosynthetic pathway that would include this compound as an

intermediate, and 1-HEP-OMe could be a breakdown product of dehydrophos rather than a biosynthetic intermediate (see Chapter 3 for a detailed discussion). Collectively, these results suggested that a positive charge on either C1 or C2 is not tolerated for DhpI activity, and that DhpI may favor an amide group and an aliphatic substituent at the α-carbon. Desmethyl dehydrophos satisfies both structural conditions, suggesting methylation might be the last step in the biosynthetic pathway (Figure 3.3).

Table 4.1 Kinetic parameters for various substrates of DhpI.

Substrates	K_m, μ**M**	k_{cat}, s^{-1}	k_{cat}/K_m, $M^{-1} \cdot s^{-1}$
DHEP	-	-	0.90
2-HEP	-	-	1.0
1-HEP	432 ± 36	(5.2 ± 0.17) x 10^{-3}	1.2 × 10^{1}
Gly-L-Leu-L-AlaP	51 ± 7	(6.9 ± 0.26) x 10^{-2}	1.4 × 10^{3}
Gly-L-Leu-D-AlaP	82 ± 4	(4.9 ± 0.068) x 10^{-2}	5.9 × 10^{2}

Desmethyl dehydrophos is not readily available using our previously described synthetic route to dehydrophos.[4] Instead, both diastereomers of the hydrogenated analog of desmethyl dehydrophos were prepared using previously described methodology[10] by Dr. Michael Kuemin. These peptides incorporate the well known phosphonate analog of alanine (AlaP; Figure 4.3). LC-MS analysis demonstrated that DhpI catalyzed the complete conversion of each of the diastereomeric peptides to their corresponding methylated derivatives (Figure 4.4*A*). The k_{cat}/K_m values of both diastereomers of Gly-L-Leu-AlaP were approximately 100-fold larger than that of 1-HEP, which displays the highest k_{cat}/K_m among the other phosphonic acids tested (Figure 4.4*B*). The enzyme did not strongly differentiate between the two diastereomers of Gly-Leu-AlaP with Gly-Leu-L-AlaP being slightly more active (2-fold higher k_{cat}/K_m) than Gly-Leu-D-AlaP. The absolute stereochemistry of one of these two diastereomers was determined by X-ray crystallography (for the crystal structure, please refer to ref.[11]). On the other hand, Gly-Leu-SerP (Figure 4.3), also prepared by Dr. Kuemin and another potential intermediate in the dehydrophos biosynthetic pathway as a precursor to the vinyl group of dehydrophos, was a very poor substrate for DhpI.

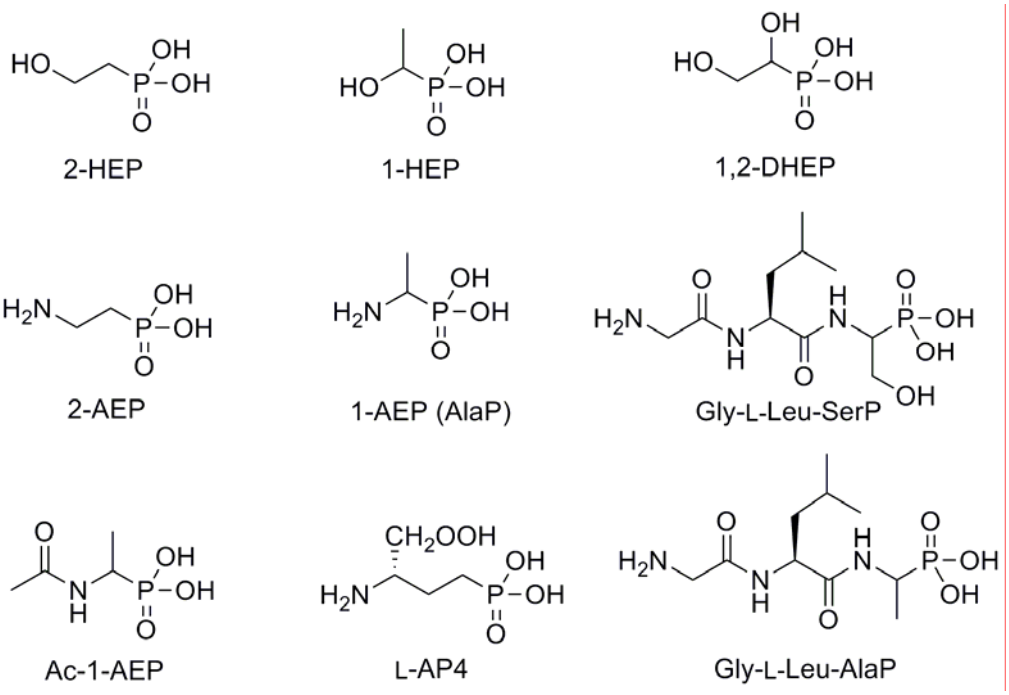


Figure 4.3 Chemical structures of various phosphonate substrate candidates used in the DhpI assays.

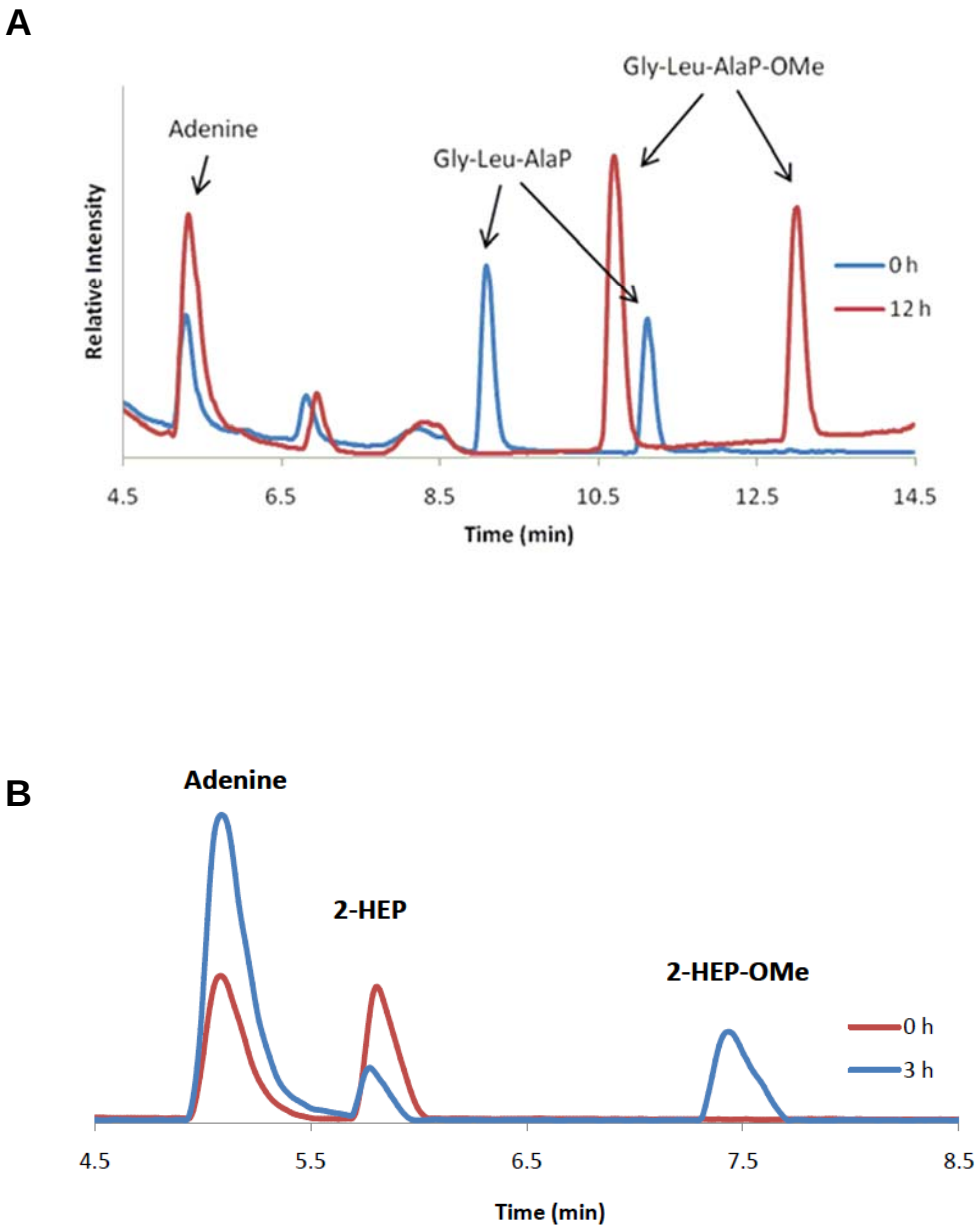


Figure 4.4 LC-MS traces showing the DhpI-catalyzed conversion of (A) Gly-L-Leu-(D,L)-AlaP to Gly-L-Leu-(D,L)-AlaP-OMe and (B) 2-HEP to 2-HEP-OMe.

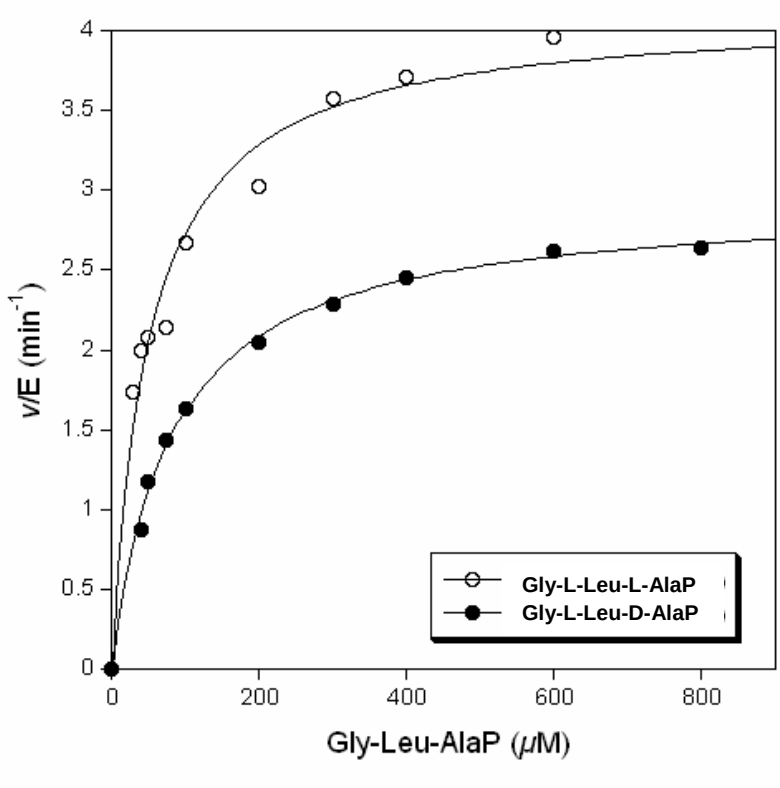


Figure 4.5 Michaelis-Menten curve obtained by using variable concentrations of Gly-Leu-L-AlaP (open circle) and Gly-Leu-D-AlaP (closed circle) at a saturated concentration of SAM (3 mM).

4.2.3. X-ray Crystal Structure of DhpI.[*]

The crystal structures of DhpI have been solved in a collaborative effort with Dr. Brian Bae in the laboratory of Prof. Satish K. Nair (Department of Biochemistry, UIUC). One structure was in complex with the substrate SAM (2.3 Å resolution, DhpI-SAM) and one in complex with the product SAH (1.5 Å resolution, DhpI-SAH). Also the ternary complex of DhpI with both SAM and a substrate analogue 2-HEP (2.3 Å resolution, DhpI-SAM-2-HEP) could be obtained. The overall structure of DhpI consists of a Rossman-fold domain alternating between seven β-strands and six α-helices similar to other methyltransferases, and the *N*-terminus of the protein is located the SAM-binding site, and the substrate binding domain is present at the *C*-terminus,.

A unique insertion is found between strand β5 and helix α6 of the Rossman-fold that covers the active site which distinguishes DhpI from the other methyltransferases. This insertion consists of a short helix and will be referred to as "capping helix" followed by two β-strands. Residues within this insertion appear to play critical roles in modulating product release following methyl transfer (Figure 4.6*A*). DhpI forms a homodimer in the presence of SAM and sulfate (and presumably with phosphonate substrates). In the DhpI-SAM-SO_4^{2-} co-crystal structure, the active site of one monomer is enclosed by the "capping helix" from an adjacent monomer, forming a domain-swapped dimer (Figure 4.6*B*). In the crystal structure of DhpI-SAM, strong electron density corresponding to a sulfate molecule can be observed adjacent to the methyl group of SAM. At the active site, several enzyme residues including Tyr15, His119, Arg168,

[*] This section on the crystallization of DhpI and the annotation of the crystal structures was performed by Dr. Brian Bae from the laboratory of Prof. Satish K. Nair (Department of Biochemistry, UIUC). More detailed discussions on the result and the methods of refinement can be found in the text and supporting information of ref.11

and Lys180 engage the sulfate anion (Figure 4.7), whose roles involved in catalysis were interrogated by mutational analysis discussed in the next section. The DhpI-SAM-SO_4^{2-} co-crystal likely mimics the binding mode of SAM with phosphonic acid substrate, prior to methyl transfer.

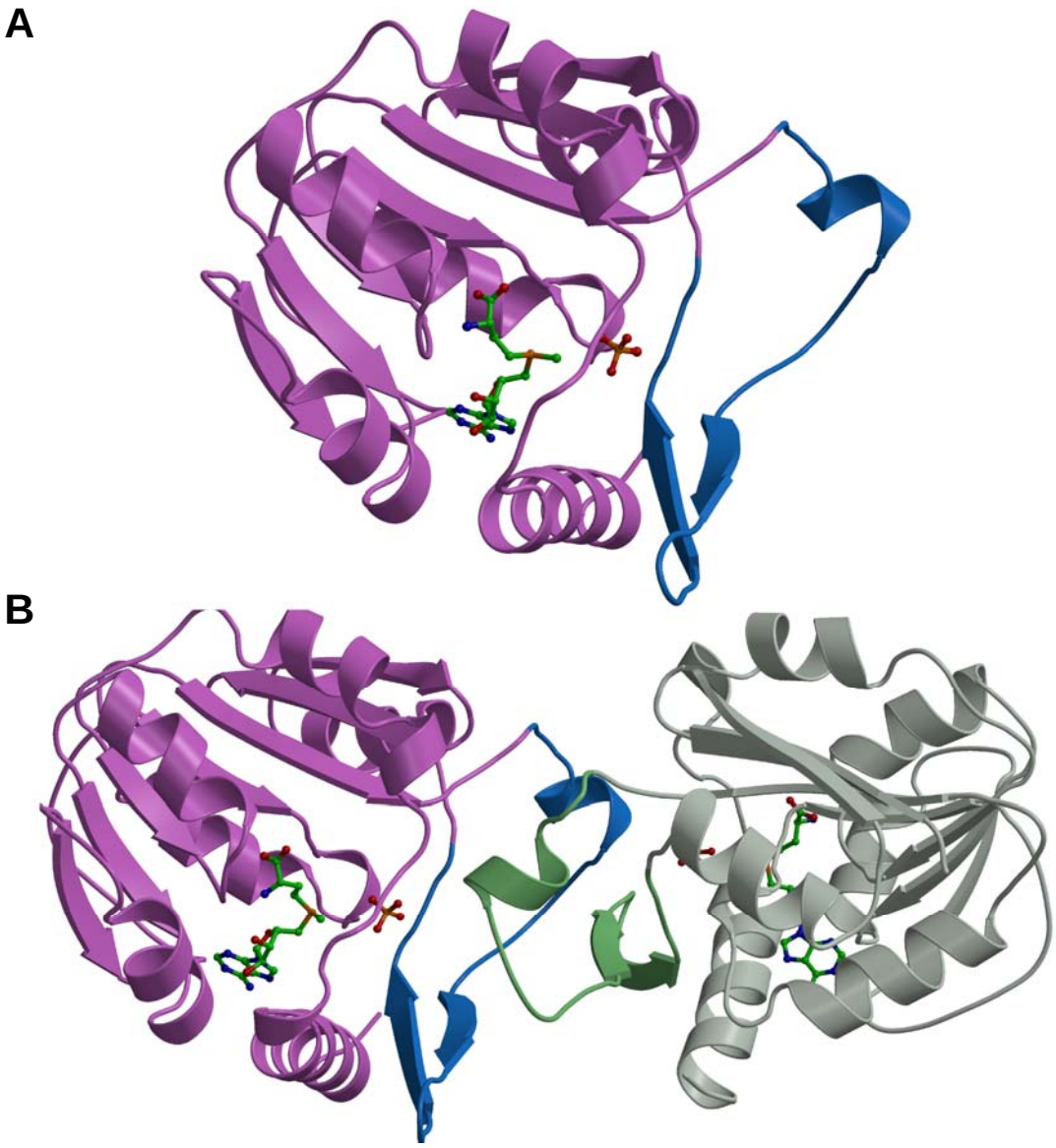


Figure 4.6 X-ray crystal structure of DhpI. (A) Co-crystal structure of DhpI with SAM and SO_4^{2-} showing a core Rossman-fold domain consisting of seven β-strands surrounded by six α-helices. A short helix and two β-strands shown in blue are referred to as the "capping helix". (B) DhpI-SAM-SO_4^{2-} co-crystal showing its homodimer formation prior to the methyl transfer. These figures were prepared by Prof. Satish K. Nair and used with his permission.

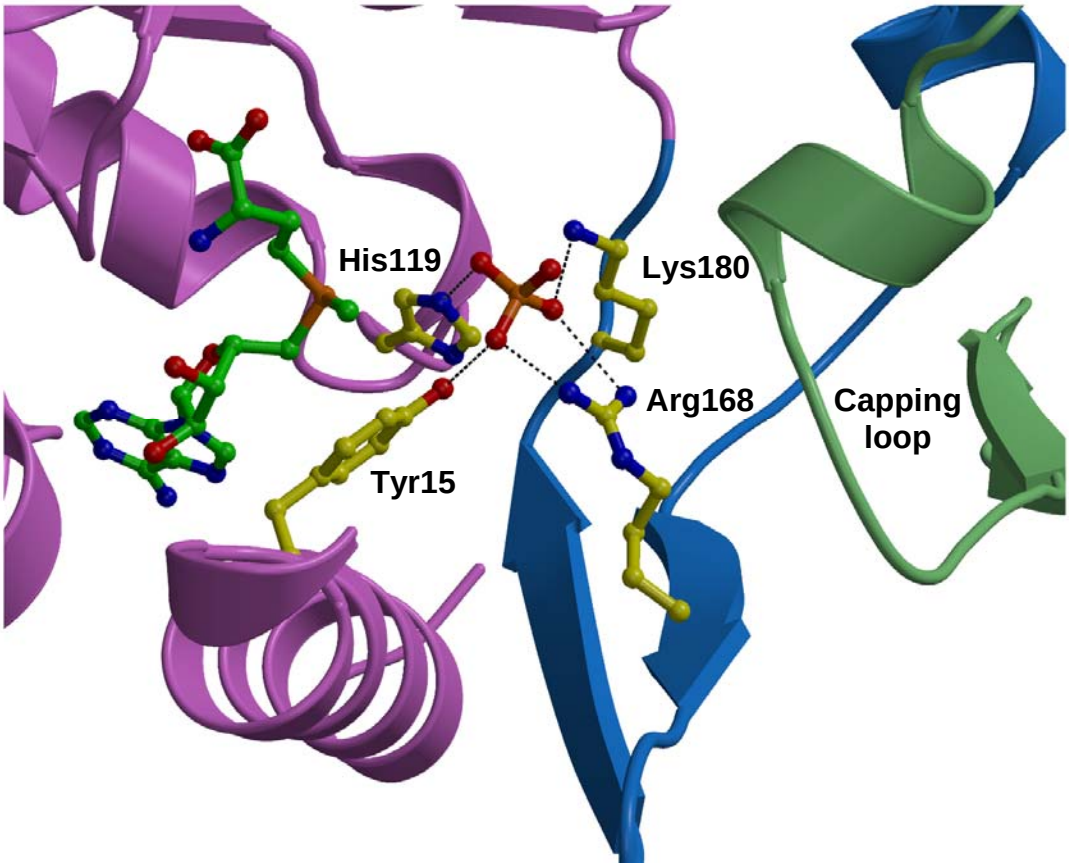


Figure 4.7 Crystal structure of DhpI with sulfate anion coordinated to four amino acid residues Tyr 15, His 119, Arg 168, and Lys 180. This figure was prepared by Prof. Satish K. Nair and used with his permission.

The overall co-crystal structure of DhpI-SAM and that of DhpI-SAH showed different crystal morphologies. DhpI-SAH is monomeric in the crystal and no electron density can be observed for the sulfate substrate to mimic in the active site. DhpI-SAM-SO_4^{2-} reveals a number of significant rearrangements near the substrate-binding site and at the "capping helix". While SAH complex forms a continuous helix (Figure 4.8*A*), in the DhpI-SAM-SO_4^{2-} structure residues Val27 through Tyr29 are rearranged to hift into the active site, resulting in a break in this helix (Figure 4.8*B*). Val27 interacts with the methyl group of SAM which facilitates the helix breakage, which is further facilitated by the presence of a proline residue ($Pro28^A$) within the helix. This helix break results in inter-subunit interactions most notably $Tyr29^A$ with $Glu155^B$ (Figure 4.9*A*). These interactions facilitate additional inter-subunit hydrogen bond interactions that results in a second notable local conformational difference between the two structures, specifically a β-hairpin with residues Val166 through Val179 moves towards the active site in the DhpI-SAM-SO_4^{2-} structure (Figure 4.8*B*), resulting in hydrogen bonding interactions between $Arg153^A$ and $Asp149^B$, the main chain oxygens of $Glu157^A$ and $Asp158^A$ with $Arg168^B$ (Figure 4.9*A*), and $Glu156^A$ with the main chain oxygen of $Gly207^B$. These interactions in the DhpI-SAM-SO_4^{2-} structure may be involved in the stabilization of the inter-subunit interactions of the "capping helix" and result in the formation of a closed active site for securing the substrate in an optimal orientation for methyl transfer (Figure 4.9*A*). Also, in DhpI-SAH complex, the "capping helix" are disordered and the continuous amino terminal helix orients $Tyr29^A$ toward the monomer, thereby disrupting interactions with $Glu155^B$ that would stabilize the "capping helix" (Figure 4.9*B*).

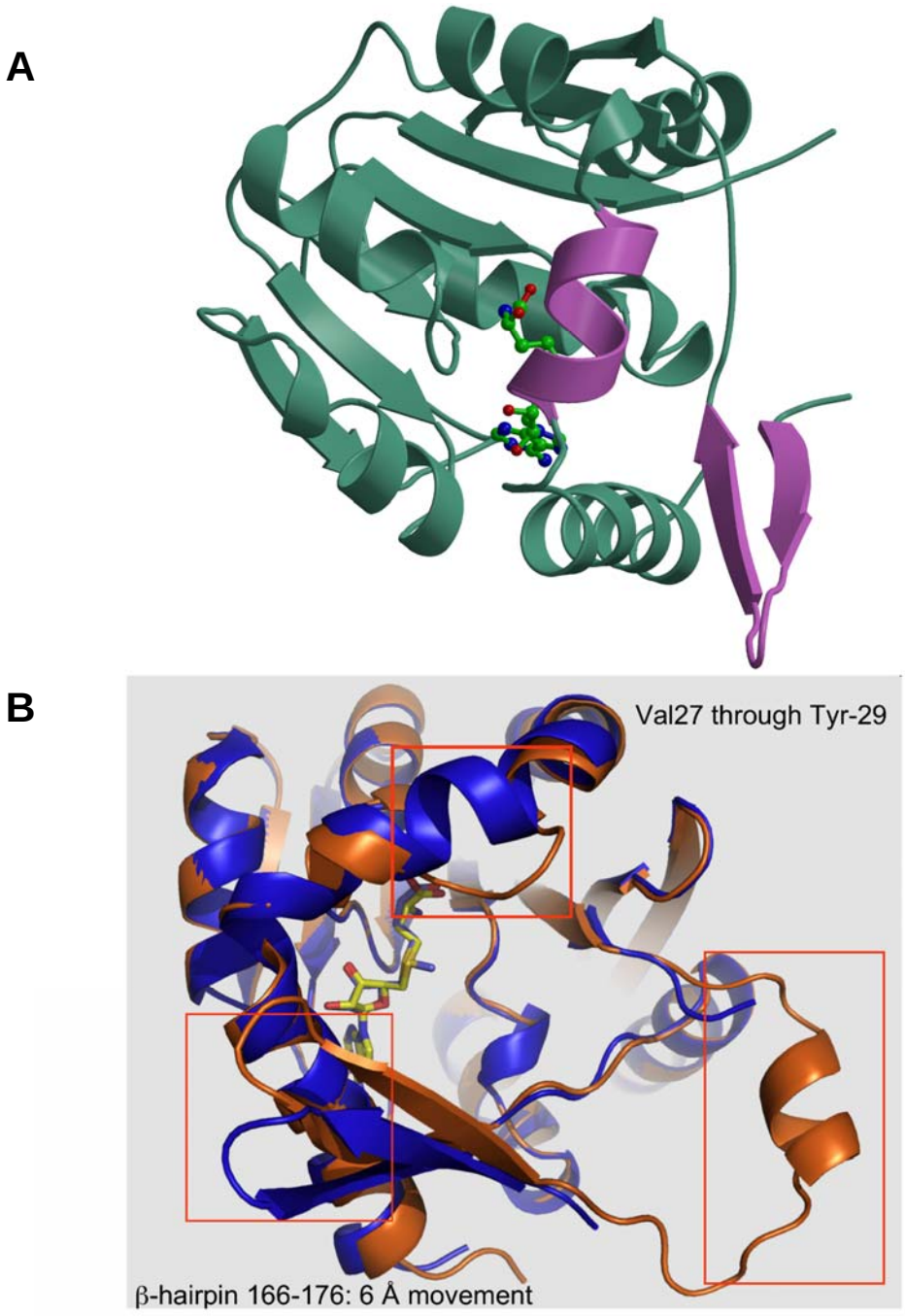


Figure 4.8 Comparison of co-crystal structures of DhpI with substrates and product. The picture was prepared by Prof. Satish K. Nair. (A) DhpI-SAH complex showing the continuous helix between His 5 and Leu 41. (B) DhpI-SAM-SO_4^{2-} structure (orange) superimposed with DhpI-SAH structure (blue) showing a rearrangement from Val 27 through Tyr 29 and β-hairpin shifting resulted from the helix break.

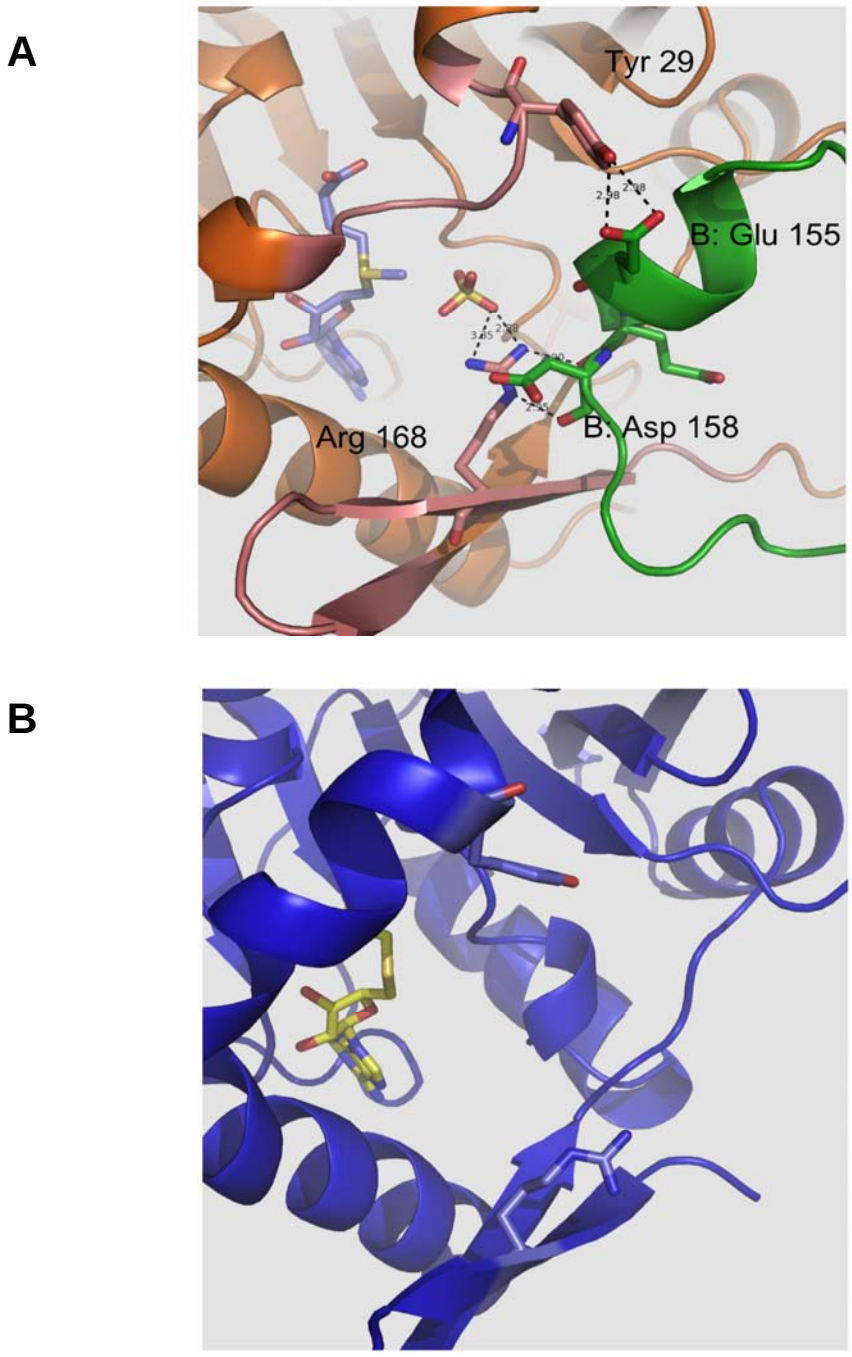


Figure 4.9 Comparison of the crystal structure of DhpI showing the inter-subunit interaction. The picture was prepared by Prof. Satish K. Nair. (A) The interactions between $Tyr29^{A}$ with $Glu155^{B}$, and $Asp158^{B}$ with $Arg168^{A}$ are shown in the DhpI-SAM-SO_4^{2-}, which result in the formation of a closed active site. (B) DhpI-SAH structure where $Tyr29^{A}$ is now oriented toward the monomer, and the interaction with $Glu155^{B}$ no longer exists.

4.2.4. Study of DhpI Active Site by Site-directed Mutagenesis.

In order to investigate the role of putative catalytic residues, several single point mutations were generated of residues that were in close proximity to the sulfate in the SAM co-crystal structure. The mutations were introduced by using Quick Change site-directed mutagenesis method. In this method DNA primers that contain the desired mutation are used in the template-based DNA amplification by polymerase chain reaction (PCR). For DhpI mutants the whole plasmid, pET15b-*dhpI* construct, was used as a template, and this wild-type plasmid, which is not mutated, could later be eliminated by using a restriction enzyme specific for methylated DNA only found in biosynthesized DNAs in contrast to unmethylated, PCR-generated DNAs. The substitution of His119, Lys180, and Arg168 with Ala resulted in loss of methylation activity with the tripeptide substrates within the detection limits of our assays. Mutation of Tyr15 to Phe did not abolish activity but did greatly decrease k_{cat} and increase K_m, resulting in overall 1,000-fold decrease in catalytic efficiency (Table 4.2). Therefore, all four residues are likely to be involved in the binding of the phosphonic acid moiety of the tripeptide substrate.

In addition, the importance of two residues that appear to be critical for the disruption of the *N*-terminal helix in the SAM co-crystal structure and for setting up the sulfate binding pocket was investigated. Mutation of Val27 that makes Vanderwaals contacts to the methyl group of SAM to Ala had a modest effect, resulting in a fivefold decrease in k_{cat}/K_m. On the other hand, mutation of Tyr29, which due to the result of helix breakage, makes a contact with Glu155 of the other subunit, had a more pronounced effect. The k_{cat}/K_m value for the tripeptide substrate decreased about 500-fold compared to the wild type enzyme.

Table 4.2 The activity of DhpI mutants as expressed in k_{cat}/K_m of Gly-L-Leu-L-AlaP.

DhpI WT and Mutants	K_m, μM[a]	k_{cat}, s^{-1}	k_{cat}/K_m, $M^{-1} \cdot s^{-1}$
WT	51 ± 7	(6.9 ± 0.26) x 10^{-2}	1.4 × 10^3
R168A	-	-	Inactive[a]
Y15F	742 ± 75	(9.4 ± 0.53) x 10^{-4}	1.3
K180A	-	-	Inactive[a]
H119A	-	-	Inactive[a]
V27A	82 ± 19	(2.0 ± 3.5) x 10^{-2}	2.4 × 10^2
Y29F	1270 ± 676	(3.6 ± 1.3) x 10^{-3}	2.8

[a] No activity was observed within our detection limits.

4.2.5. Applications of DhpI in the Esterification of Phosphonic acids

Given the improved pharmacokinetic properties of phosphonate esters compared to phosphonic acids, investigation of the substrate scope of DhpI was pursued. Encouraged by the observed broad substrate scope of DhpI, the enzymatic esterification of other naturally synthesized phosphonic acids which are known to exhibit interesting biological activities was tested. The enzyme assays were performed in a similar fashion as described previously, with LC-MS as the method of product detection. The enzyme was able to fully convert the anti-malaria clinical candidate fosmidomycin to its methyl ester (Figure 4.10*A*, 4.11*A*). Extracted ion chromatograms show the complete methylation of fosmidomycin within 24 h. Fosmidomycin, which has a retention time (RT) of 5.9 min (m/z 184) was not present, and instead a new peak which corresponds to monomethylated fosmidomycin (m/z 198) was observed with a RT of 9.4 min. Similarly, the clinically used antibiotic fosfomycin (RT = 6.1 min, m/z 139) was completely transformed to its methyl ester (RT = 9.1 min, m/z 153) by DhpI in less than 24 h (Figure 4.10*B*, 4.11*B*). The substrate promiscuity of the enzyme was also extended to L-(+)-2-amino-4-phosphonobutyric acid (L-AP4, Figure 4.3), an often used non-hydrolizable analog of phosphoSer. Another interesting phosphonate that could be esterified by DhpI is methyl phosphonate. Collectively, these experiments demonstrate the promise of using DhpI or in vitro evolved analogs for bioengineering of esterified phosphonates.

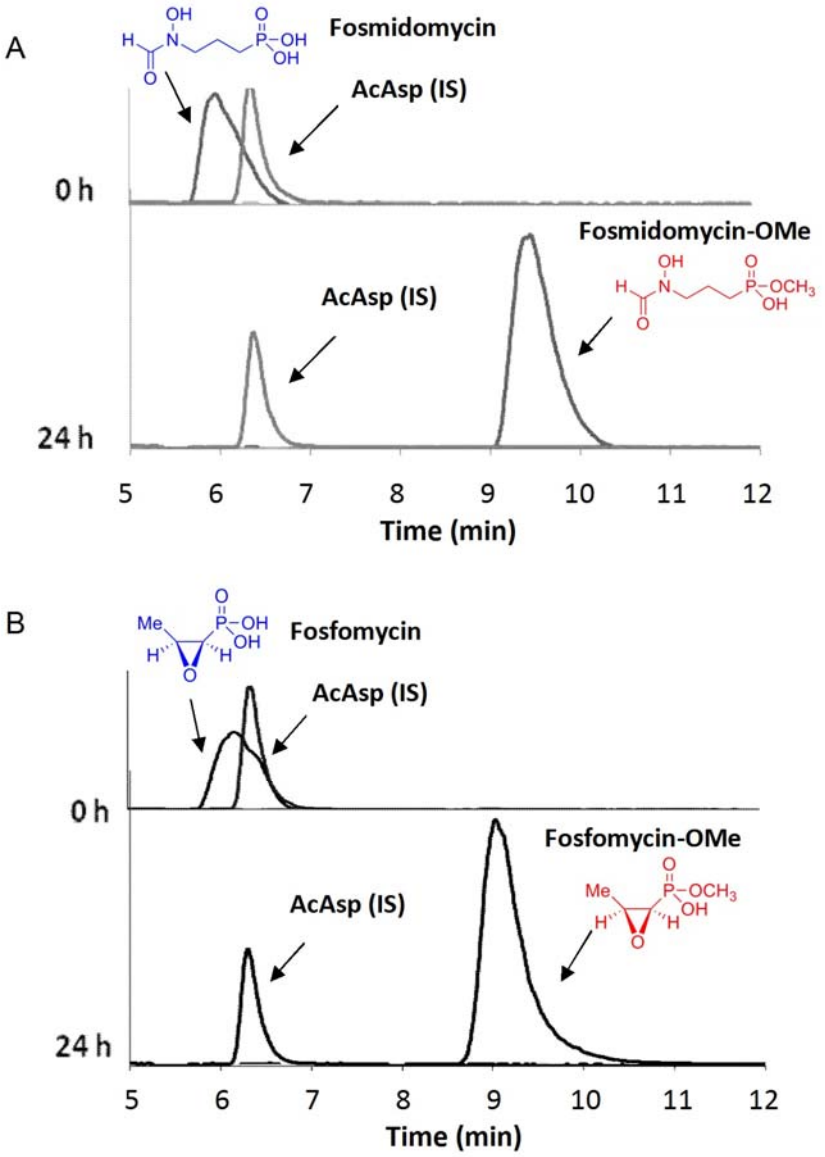


Figure 4.10 Methylation of phosphonate natural products. Extracted ion chromatograms showing the complete methylation of fosmidomycin and fosfomycin as substrate of DhpI after 24 h of incubation. (*A*) Fosmidomycin (RT = 5.9 min, m/z 184) was fully converted to fosmidomycin-OMe (RT = 9.4 min, m/z 198). (*B*) Fosfomycin (RT = 6.1 min, m/z 139) was fully converted to fosfomycin-OMe (RT = 9.1 min, m/z 153). IS = internal standard.

A

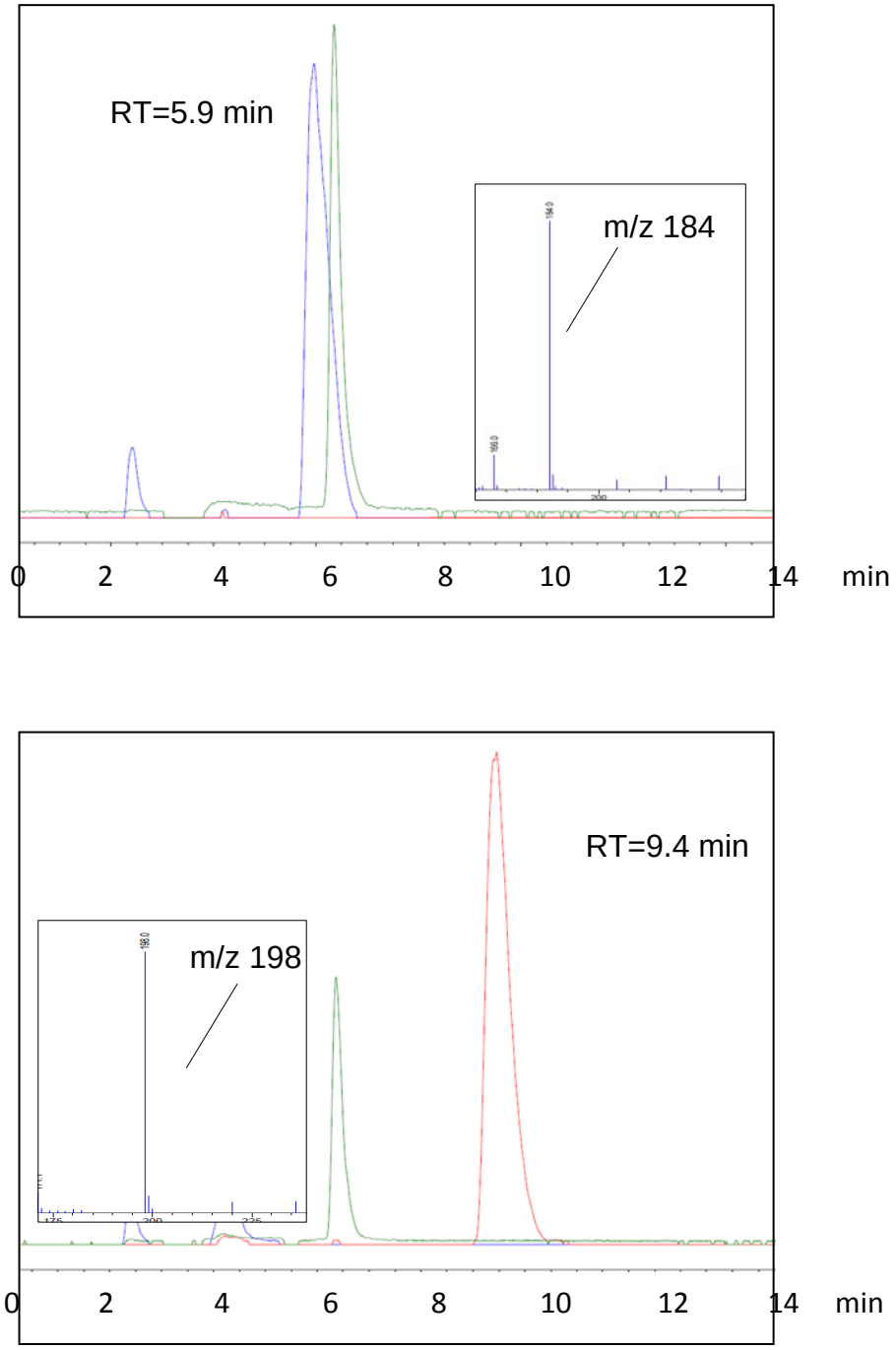


Figure 4.11 Mass spectra showing the methylation of fosmidomycin and fosfomycin by DhpI. Extracted ion chromatograms are shown that correspond to $(M+H)^+$ of the substrate (blue), product (red), and the internal standard (IS), AcAsp (green). Insets show the mass spectrum at the retention times (RT) shown. (A) Fosmidomycin (RT = 5.9 min, m/z 184) was fully converted to fosmidomycin-OMe (RT = 9.4 min, m/z 198). (B) Fosfomycin (RT=6.1 min, m/z 139) was fully converted to fosfomycin-OMe (RT=9.1 min, m/z 153).

B

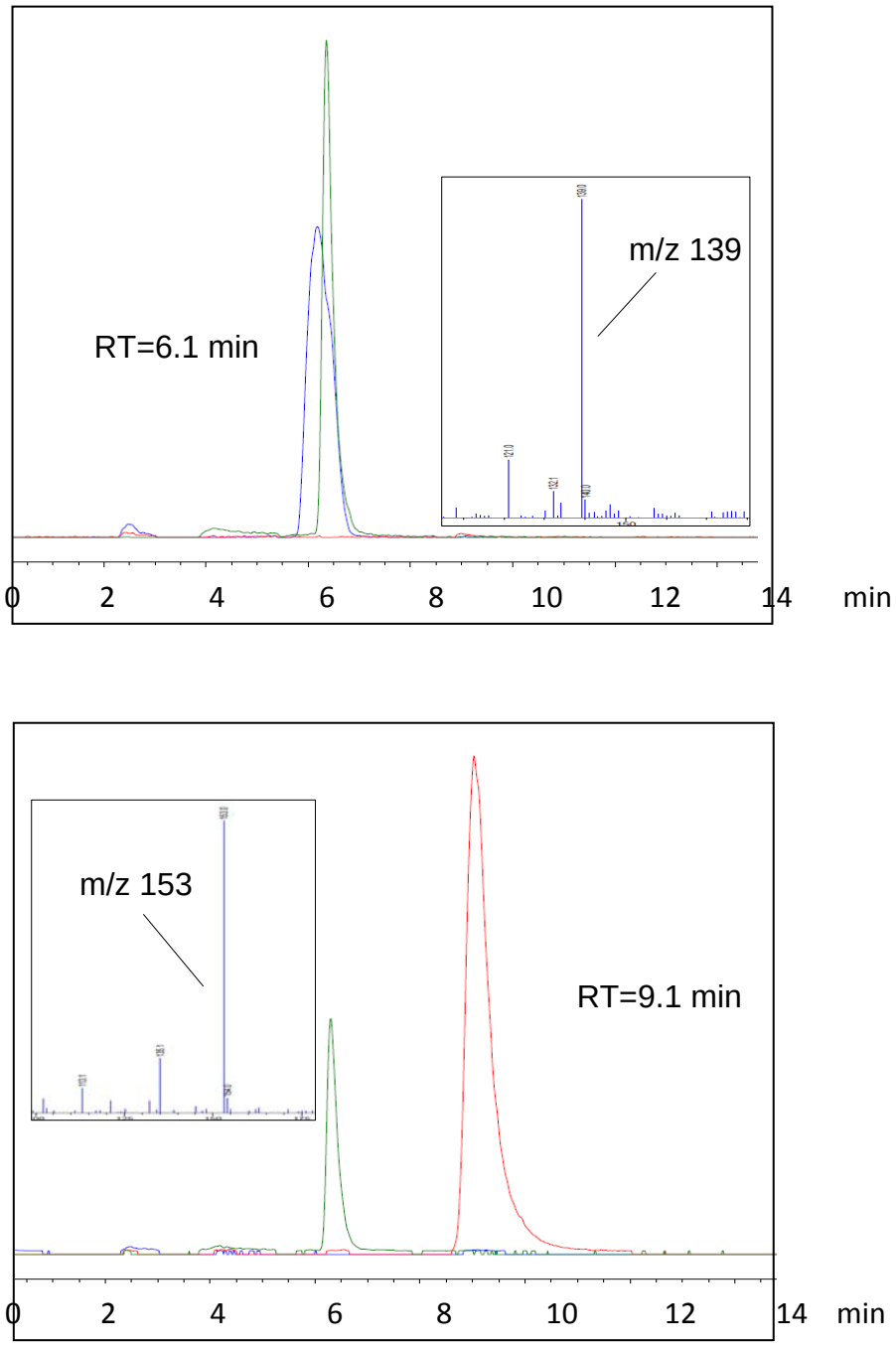


Figure 4.11 (continued)

The application of DhpI for alkylating phosphonic acids was further extended to utilization of the SAM analogue allyl-AdoMet (Figure 4.12), a SAM analogue prepared by Dr. Kuemin using chemical synthesis. In the LC-MS analysis of the assay, partial conversion of the phosphonate substrate, Gly-Leu-L-AlaP (RT = 9 min, m/z 296) to Gly-Leu-L-AlaP-OAllyl (RT = 15 min, m/z 336) could be observed (Figure 4.13). This result, although quite preliminary, demonstrates the applicability of allyl-AdoMet as alkyl donor in DhpI-catalyzed alkylation reaction.

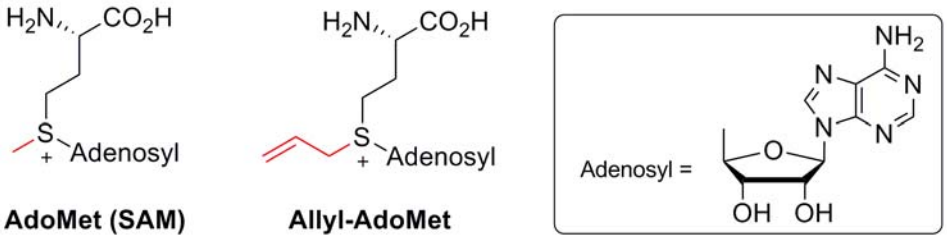


Figure 4.12 Structure of AdoMet and its allyl analogue "allyl-AdoMet".

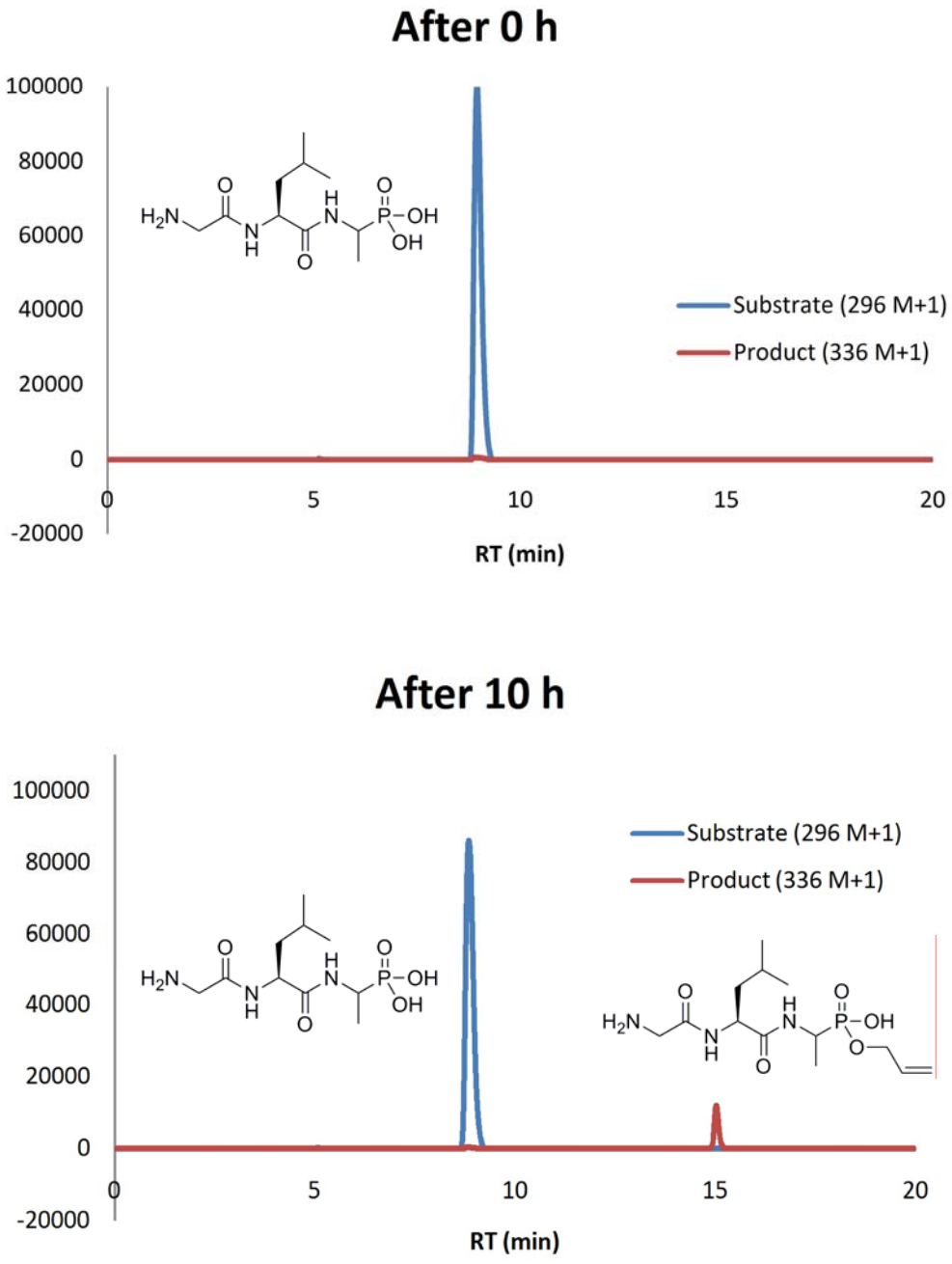


Figure 4.13 Formation of monoallyl phosphonate ester, Gly-Leu-AlaP-OAllyl. Shown are the extracted ion chromatograms for the substrate (Gly-Leu-AlaP, M+1 = 296, blue chromatograms) with RT = 9 min and the proposed product (Gly-Leu-AlaP-OAllyl, M+1 = 336, red chromatograms) with RT = 15 min. The samples were taken at 0 h and 10 h after the initiation of the reaction.

4.3 – Discussion and Conclusion

PEP mutase installs the P-C bond in all phosphonates for which the gene clusters are currently known. A multidisciplinary team focused on naturally occurring phosphonates has recently demonstrated that the gene clusters of these compounds can be readily identified using the *pepM* gene as a marker.[12-21] In addition, these studies have shown that phosphonate biosynthesis is widespread and that new clusters that encode for as yet structurally unidentified phosphonates are readily found.[22] Given the range of biological activities that known phosphonates display and given their current applications in medicine and agriculture, these unidentified phosphonates may provide a reservoir for new compounds with useful activities. As noted, the bisanionic charge state of phosphonates provides a potential hurdle for their use, a drawback that in synthetic phosphonates is circumvented by esterification. Dehydrophos is currently the only known esterified naturally occurring phosphonate and DhpI is the only SAM-dependent *O*-methyltransferase that transfers the methyl group to a phosphonate moiety.

In this study, we identified tripeptides closely resembling desmethyl dehydrophos as the preferred substrates of the enzyme, strongly suggesting it acts as the last step in the biosynthetic pathway to dehydrophos. This observation argues against one of the suggested roles of the methyl ester: as a protective measure to prevent the biosynthetic intermediates from inhibiting endogenous enzymatic reactions in the producing strain. Instead, introduction of the methyl group in the last step suggests that the methyl group is important for biological activity on other surrounding organisms. Acetylphosphonate methyl ester, the expected breakdown product after hydrolysis of the amide bonds of dehydrophos within the targeted organism, is a known inhibitor of pyruvate oxidase and

the esterified compound is a 125-fold better inhibitor than acetylphosphonate itself.[7] Thus, our current hypothesis for methylation is to provide a better mimic of the monoanionic pyruvate substrate. Further analysis with both diastereomers of the Gly-Leu-SerP tripeptide did not display good activity in DhpI catalyzed methylation reaction, eliminating the possibility that the methylation takes place prior to phosphorylation and elimination. Instead, the hydroxyl group on the beta carbon in Gly-Leu-SerP may create a steric and/or electronic perturbation in the active site of DhpI, as all the other substrates with high conversion rate contained hydrophobic residues. Otherwise, DhpI displayed broad substrate specificity and methylated two clinically important phosphonate natural products. As such, it may prove to be a very useful biocatalyst to prepare methylated analogs of naturally occurring phosphonates by fermentation in engineered organisms. In addition, this study shows that the SAM analogue allyl-AdoMet can also be used to alkylate phosphonates using DhpI. Not only does this observation suggest broad utility of DhpI as an alkylating agent for phosphonate compounds, but this work also sets a stage for further exploitation of DhpI as a tool for the identification of unknown phosphonates as well as for their purification from cellular metabolites by using keto-AdoMet as methyl donor and subsequent use of the installed keto group as a handle in bioorthogonal reactions for labeling or tagging the unknown compounds.

In addition to providing important new information regarding the biosynthesis of dehydrophos and showing promise for use as a biocatalyst, our work also uncovered unique mechanistic insights into DhpI catalysis. The three co-crystal structures illustrate an unusual conformational change that is triggered by the presence of both the methyl group of SAM and bound sulfate/phosphonate. As both the SAM and SAH co-crystals

were grown under identical conditions (i.e. equivalent sulfate concentrations), the presence of the anion in only the SAM complex argues that the presence of the methyl group instigates the structural rearrangements that are further stabilized by substrate binding. The movement of Val27 in towards the active site in the DhpI-SAM-SO_4^{2-} structure provides a solvent-excluded binding pocket for the methyl group. This movement disrupts the continuous helix containing Val27 that is seen when SAH is present and is likely aided by the presence of Pro28. The energetic penalty for disrupting the helix is compensated by four new intersubunit interactions that stabilize the capping helix. Mutation of Val27 to Ala resulted in only a five-fold decrease in the catalytic efficiency over the wild-type, indicating that the Vanderwaals interaction between Val27 and the methyl group of SAM provides only a limited energetic driving force for the structural rearrangement. It is likely that additional energy may be provided by the electrostatic interaction between the negatively charged sulfate/phosphonate group and the positively charged sulfonium group of SAM. In the SAH structure, these interactions are likely destabilizing, favoring the continuous helix and disrupting the extensive intersubunit interactions and preventing the stabilization of the sulfate/phosphonate binding pocket.

4.4 - Experimental Procedures

4.4.1. Preparation of DhpI Mutants

General Materials and Methods for Site-directed Mutagenesis. All molecular biology techniques, including PCR, transformation and plasmid preparation, were performed using standard procedures.[23] The amino acid mutations were introduced by the Quick-Change method as described in the Stratagene manufacturer's guide. DpnI was purchased from Invitrogen (Carlsbad, CA), and the oligonucleotides used for mutations were purchased from IDT. PCR amplifications were performed using an automated thermocycler (PTC 150, MJ Research). Plasmid isolation was accomplished by Qiagen Miniprep kits (Valencia, CA). Chemically competent *E. coli* DH5α cells were purchased from the University of Illinois at Urbana-Champaign (UIUC) Cell Media Facility. Sequencing reactions were performed at the W.M. Keck Center for Biotechnology at the University of Illinois at Urbana-Champaign utilizing standard T7 forward and reverse primers (T7-for/T7-rev). Chemically competent *E. coli* DH5α cells were transformed with each ligation mixture and plated on Luria-Bertani (LB)-agar containing the appropriate antibiotic to select for positive clones. For pET15b constructs ampicillin (Sigma-Aldrich) was used at 100 μg/mL concentration.

Preparations of pET15b-dhpI (R168A) [JHL-VII-81, 89], pET15b-dhpI (Y15F) [JHL-IX-6], pET15b-dhpI (K180A) [JHL-IX-8], pET15b-dhpI (H119A) [JHL-IX-8]. The construct pET15b-dhpI was obtained from Benjamin T. Circello from the laboratory of Prof. William W. Metcalf (Department of Microbiology, UIUC). Single point mutations were introduced by the QuickChange site-directed mutagenesis method using Pfu DNA polymerase and pET15b-dhpI wild type DNA as template for PCR. The oligonucleotide

primers were designed to contain mutation coding gene sequences (Table 4.3). Wild type DNA in the pET15b vector, which was isolated from *E. coli* was dam methylated and therefore could be digested by DpnI endonucleases leaving only DNA obtained by PCR in circular plasmid form. The *E. coli* DH5α strain was transformed with the DNA that contained the desired mutation.

Table 4.3 Oligonucleotides used in the QuickChange site-directed mutagenesis of DhpI.

Oligonucleotide	DNA sequence
DhpI WT FOR	5'-GGATCCGAAGGAAAGCTTCATATGACCACTTCGCACGG-3'
DhpI WT REV	5'-GGCGCGCCAGATCTGAATTCTCACCGCGGTCCGGGGCGGCACG-3'
DhpI R168A FOR	5'-CCCGAGGTCGCCGTCCGCGCCACCCTCCAGGACGGGCGC-3'
DhpI R168A REV	5'-GCGCCCGTCCTGGAGGGTGGCGCGGACGGCGACCTCGGG-3'
DhpI Y15F FOR	5'-GGAGGCCCGGGCCCGGAAGTAGGACAGCTGGCTCTCGATCAGACC-3'
DhpI Y15F REV	5'-GAGAGCCAGCTGTCCTACTTCCGGGCCCGGGCCTCC-3'
DhpI K180A FOR	5'-CGCAGCTTCCGCATCGTGGCGGTCTTCCGCAGCCCGGCC-3'
DhpI K180A REV	5'-GGCCGGGCTGCGGAAGACCGCCACGATGCGGAAGCTGCG-3'
DhpI H119A FOR	5'-GCCCACTGGCTGGCCGCTGTGCCCGACGACCGG-3'
DhpI H119A REV	5'-CCGGTCGTCGGGCACAGCGGCCAGCCAGTGGGC-3'
DhpI V27A FOR	5'-TACGACGCGACCTTCGCGCCGTACATGGACTCC-3'
DhpI V27A REV	5'-GGAGTCCATGTACGGCGCGAAGGTCGCGTCGTA-3'
DhpI Y29F FOR	5'-GCGACCTTCGTGCCGTTCATGGACTCCGCGGCG-3'
DhpI Y29F REV	5'-CGCCGCGGAGTCCATGAACGGCACGAAGGTCGC-3'

4.4.2. Expression and Purification of DhpI and SAH nucleosidase

E. coli Rosetta2 (DE3) cells harboring the DhpI-encoding plasmid were grown in a Luria-Bertani (LB) medium supplemented with antibiotic markers (100 μg/mL ampicillin and 20 μg/mL chloramphenicol) at 37 °C until OD_{600} reached 0.5 - 0.7. The expression of *N*-terminally His_6-tagged DhpI was induced by addition of isopropyl-1-thio-β-D-galactoside (IPTG) to 0.5 mM and the cells were continued to grow at 18 °C for 15 h. The protein was purified by nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography (Qiagen, Valencia, CA). In a typical purification of the expressed proteins, *E. coli* cells from a 3 L culture were harvested by centrifugation at 5,000 × *g* for 15 min at 4 °C. The cell pellet was resuspended in 30 mL of ice cold buffer A (50 mM Tris·HCl, pH 7.5, 100 mM NaCl, 10 mM imidazole) supplemented with 0.4 mg/mL lysozyme and protease inhibitor cocktail tablets (Complete Mini, EDTA-free, Roche), and cells were disrupted in a chilled French press cell at 15,000 psi. The cell lysate was cleared by centrifugation at 25,000 × *g* for 30 min at 4 °C, and the cleared soluble fraction was loaded on a column containing 8 mL Ni-NTA affinity resin (Qiagen, Valencia, CA) equilibrated with buffer A. After the loaded sample was loaded onto the column, unbound proteins were washed from the column with 10 column volumes of buffer B (50 mM Tris·HCl, pH 7.5, 100 mM NaCl, 20 mM imidazole) to remove any proteins nonspecifically bound to the column. Elution of the protein was achieved with buffer C (50 mM Tris·HCl, pH 7.5, 100 mM NaCl, 50-500 mM imidazole) with increasing imidazole concentrations in a step gradient from 50 mM to 500 mM with 30 mM increments with a total elution volume of 200 mL. The elution fractions containing the desired protein as detected by SDS-PAGE gels stained with Coomassie Brilliant Blue

were concentrated by ultrafiltration using an Amicon Ultra 10 kDa membrane (Millipore). Imidazole and excess salt were removed by passage over a PD-10 column (GE Healthcare) pre-packed with Sephadex G-25. The protein was eluted from the column with buffer D (50 mM Tris·HCl, pH 7.5, 100 mM NaCl, 20% (w/v) glycerol), and the purified protein was flash frozen and stored at – 80 °C. Protein concentration was determined by Bradford assay following the standard protocol using bovine serum albumin as a standard protein. The overall yield of >95% pure protein obtained ranged between 10-15 mg per liter of culture.[23] For the SAH nucleosidase, the DNA construct which encodes for *N*-terminally hexahistidine-tagged protein was generously provided by Prof. John E. Cronan (Department of Microbiology, UIUC), and the overexpression and purification was carried out in similar fashion as described for DhpI.

4.4.3. Preparation of Substrates and DhpI Activity Assay

Syntheses of 2-HEP, 1-HEP, 1,2-DHEP, Ac-1-AEP, and their monomethylated phosphonate esters were carried out by myself, and are outlined in Chapter 3. Compounds 1-AEP, 2-AEP, and L-AP4 were commercially available. Syntheses of the tripeptide Gly-Leu-AlaP, its methylated product and Gly-Leu-SerP were achieved by a postdoctoral fellow, Dr. Michael Kuemin who also prepared the SAM analogues including Allyl-AdoMet (See reference 11 for details of the synthesis of the tripeptide).

Enzyme kinetic assays were carried out at 30 °C in 50 mM Tris·HCl (pH 7.8) in 1.5 mL reaction vials. The reaction mixture contained 3 mM AdoMet, 2 - 24 μM DhpI, 1 μM AdoHcy nucleosidase, and 30 – 2000 μM phosphonate substrates in a final volume of 300 μL. After adding DhpI, the reaction mixture was preincubated at 30 °C for 3 min before the reaction was initiated by addition of AdoMet. For quantification, 30 μL aliquots of the assay mixture were removed at designated time points and quenched with an equal volume of 0.1% formic acid containing 200 μM acetylated aspartate (AcAsp) as internal standard (IS) for quantification during LC-MS analysis. The assay sample was injected to a Synergi C18 RP-fusion column (150 mm × 4.6 mm, Phenomenex) on a HPLC system (Agilent Technologies 1200 series) equipped with a multimode electrospray ionization/APCI spray chamber. HPLC parameters were as follows: column temperature of 25 °C; solvent A, 0.1% formic acid in water; solvent B, methanol; for non-peptide phosphonates: isocratic mobile phase of 100% solvent A, for tripeptide-like phosphonates: gradient from 100% A to 30% B over 13 min; flow rate, 0.5 ml/min; detection by APCI/MS operating in a positive-ion mode.

4.4.4. MS Data Processing and Kinetic Analysis

For each sample, the eluting fractions from HPLC were ionized by APCI and the mass spectrum was detected in positive ion mode. Relative intensities were determined by integrating the selected ion chromatogram for a specific ion across the entire elution profile. Quantification of each analyte was achieved by constructing a standard curve of chemically synthesized authentic sample. Each signal was normalized with a response factor (*i.e*. ratio of the integrated area of the analyte peak to the integrated area of the internal standard, AcAsp) to account for variations arising from the instrument and sampling conditions. The initial rates of the enzyme reactions were plotted for at least eight varying concentrations of substrates, and fitted to the Michaelis-Menten equation using nonlinear regression (KaleidaGraph program, ver.3.5) to estimate the apparent steady-state rate constants. All of the assays used for determination of k_{cat} and K_m were performed in duplicate.

4.5 – References

1. Tachibana, K.; Watanabe, T.; Sekizawa, Y.; Takematsu, T.: "Action mechanism of bialaphos .1. Inhibition of glutamine-synthetase and quantitative changes of free amino acids in shoots of bialaphos-treated Japanese barnyard millet." *J. Pestic. Sci.* **1986**, *11*, 27-31.

2. Leason, M.; Cunliffe, D.; Parkin, D.; Lea, P. J.; Miflin, B. J.: "Inhibition of pea leaf glutamine-synthetase by methionine sulfoximine, phosphinothricin and other glutamate analogs." *Phytochemistry* **1982**, *21*, 855-857.

3. Hecker, S. J.; Erion, M. D.: "Prodrugs of phosphates and phosphonates." *J. Med. Chem.* **2008**, *51*, 2328-2345.

4. Whitteck, J. T.; Ni, W.; Griffin, B. M.; Eliot, A. C.; Thomas, P. M.; Kelleher, N. L.; Metcalf, W. W.; van der Donk, W. A.: "Reassignment of the structure of the antibiotic A53868 reveals an unusual amino dehydrophosphonic acid." *Angew. Chem. Int. Ed.* **2007**, *46*, 9089-9092.

5. Allen, J. G.; Havas, L.; Leicht, E.; Lenoxsmith, I.; Nisbet, L. J.: "Phosphonopeptides as antibacterial agents - metabolism and pharmacokinetics of Alafosfalin in animals and humans." *Antimicrob. Agents Chemother.* **1979**, *16*, 306-313.

6. Atherton, F. R.; Hall, M. J.; Hassall, C. H.; Lambert, R. W.; Lloyd, W. J.; Lord, A. V.; Ringrose, P. S.; Westmacott, D.: "Phosphonopeptides as substrates for peptide-transport systems and peptidases of *Escherichia coli*." *Antimicrob. Agents Chemother.* **1983**, *24*, 552-528.

7. Obrien, T. A.; Kluger, R.; Pike, D. C.; Gennis, R. B.: "Phosphonate analogs of pyruvate: Probes of substrate binding to pyruvate oxidase and other thiamin pyrophosphate-dependent decarboxylases." *Biochim. Biophys. Acta* **1980**, *613*, 10-17.

8. Circello, B. T.; Eliot, A. C.; Lee, J.-H.; van der Donk, W. A.; Metcalf, W. W.: "Molecular cloning and heterologous expression of the dehydrophos biosynthetic gene cluster." *Chem. Biol.* **2010**, *17*, 402-411.

9. Hendricks, C. L.; Ross, J. R.; Pichersky, E.; Noel, J. P.; Zhou, Z. S.: "An enzyme-coupled colorimetric assay for *S*-adenosylmethionine-dependent methyltransferases." *Anal. Biochem.* **2004**, *326*, 100-105.

10. Soroka, M.; Zygmunt, J.: "Tritylamine (triphenylmethylamine) in organic synthesis .1. The Synthesis of *N*-(triphenylmethyl)Alkanimines, 1-(triphenylmethylamino)alkylphosphonic esters, and 1-aminoalkylphosphonic acids and esters." *Synthesis* **1988**, 370-375.

11. Lee, J.-H.; Bae, B.; Circello, B. T.; Metcalf, W. W.; Nair, S. K.; van der Donk, W. A.: "Characterization and structure of DhpI, a novel phosphonate *O*-methyltransferase involved in dehydrophos biosynthesis." *Submitted for publication* **2010**.

12. Seto, H.; Kuzuyama, T.: "Bioactive natural products with carbon-phosphorus bonds and their biosynthesis." *Nat. Prod. Rep.* **1999**, *16*, 589-596.

13. Woodyer, R. D.; Shao, Z. Y.; Thomas, P. M.; Kelleher, N. L.; Blodgett, J. A. V.; Metcalf, W. W.; Van der Donk, W. A.; Zhao, H. M.: "Heterologous production of fosfomycin and identification of the minimal biosynthetic gene cluster." *Chem. Biol.* **2006**, *13*, 1171-1182.

14. Woodyer, R. D.; Li, G. Y.; Zhao, H. M.; van der Donk, W. A.: "New insight into the mechanism of methyl transfer during the biosynthesis of fosfomycin." *Chem. Commun.* **2007**, 359-361.

15. Eliot, A. C.; Griffin, B. M.; Thomas, P. M.; Johannes, T. W.; Kelleher, N. L.; Zhao, H. M.; Metcalf, W. W.: "Cloning, expression, and biochemical characterization of *Streptomyces rubellomurinus* genes required for biosynthesis of antimalarial compound FR900098." *Chem. Biol.* **2008**, *15*, 765-770.

16. Johannes, T. W.; DeSieno, M. A.; Griffin, B. M.; Thomas, P. M.; Kelleher, N. L.; Metcalf, W. W.; Zhao, H. M.: "Deciphering the late biosynthetic steps of antimalarial compound FR-900098." *Chem. Biol.* **2010**, *17*, 57-64.

17. Borisova, S. A.; Circello, B. T.; Zhang, J. K.; van der Donk, W. A.; Metcalf, W. W.: "Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by *Bacillus subtilis* ATCC6633." *Chem. Biol.* **2010**, *17*, 28-37.

18. Blodgett, J. A. V.; Thomas, P. M.; Li, G. Y.; Velasquez, J. E.; van der Donk, W. A.; Kelleher, N. L.; Metcalf, W. W.: "Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide." *Nat. Chem. Biol.* **2007**, *3*, 480-485.

19. Cicchillo, R. M.; Zhang, H. J.; Blodgett, J. A. V.; Whitteck, J. T.; Li, G. Y.; Nair, S. K.; van der Donk, W. A.; Metcalf, W. W.: "An unusual carbon-carbon bond cleavage reaction during phosphinothricin biosynthesis." *Nature* **2009**, *459*, 871-U810.

20. Shao, Z. Y.; Blodgett, J. A. V.; Circello, B. T.; Eliot, A. C.; Woodyer, R.; Li, G. Y.; van der Donk, W. A.; Metcalf, W. W.; Zhao, H. M.: "Biosynthesis of 2-hydroxyethylphosphonate, an unexpected intermediate common to multiple phosphonate biosynthetic pathways." *J. Biol. Chem.* **2008**, *283*, 23161-23168.

21. Lee, J. H.; Evans, B. S.; Li, G. Y.; Kelleher, N. L.; van der Donk, W. A.: "*In vitro* characterization of a heterologously expressed nonribosomal peptide synthetase involved in phosphinothricin tripeptide biosynthesis." *Biochemistry* **2009**, *48*, 5054-5056.

22. Metcalf, W. W.; van der Donk, W. A.: "Biosynthesis of phosphonic and phosphinic acid natural products." *Annu. Rev. Biochem.* **2009**, *78*, 65-94.

23. Sambrook, J. F., E. F.; Maniatis, T.: *Molecular cloning: a laboratory manual* edn 2nd ed. New York: Cold Spring Harbor Laboratory: Cold Spring Harbor; 1989.

AUTHOR'S BIOGRAPHY

Jin Hee Lee graduated from Hanyang University at Seoul, Korea in 2001 with a B.S. degree in Chemistry. She continued her study at Hanyang University and received a M.S. degree in Organic Chemistry in 2003 during which she worked on studies of the Pd-catalyzed reactions of 3,5-dibromo-2-pyrone under the supervision of Professor Cheon-Gyu Cho. After working for a year in the Synthetic Laboratory for Bioactive Molecules as a research assistant, she came to study at the Chemistry Department of the University of Illinois at Urbana-Champaign and subsequently joined Professor Wilfred van der Donk's research group in 2004 where she worked on the investigation of the biosynthesis of the phosphonate antibiotics bialaphos and dehydrophos. Upon completing her doctoral degree in 2010, she will begin her post-doctoral work with Professor Jon Thorson at the University of Wisconsin at Madison.